

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 12:24:31; Search time 165 Seconds

(without alignments)
35.160 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDLRAYGFWRLL 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database: A_Geneseq_16Dec04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	100.0	15	2	Aay30352 Epitope d
2	87	100.0	15	5	Aae19240 Streptoco
3	87	100.0	66	5	Aae19245 Bi-peptid
4	82	94.3	15	5	Aae19244 Streptoco
5	82	94.3	47	5	Aae19247 Tri-pepti
6	47	54.0	554	4	Abg19785 Novel hum
7	47	54.0	554	4	Abg19731 Novel hum
8	47	54.0	1007	4	Abg03083 Novel hum
9	45	51.7	741	6	Abp79868 N. gonorr
10	45	51.7	903	7	Adf03901 Bacterial
11	44	50.6	118	5	Abp02252 Human ORF
12	44	50.6	267	4	Abg26189 Novel hum
13	44	50.6	1096	5	Aae25282 Human nuc
14	44	50.6	1114	3	Aab08517 Amino aci
15	44	50.6	1262	4	Aam78463 Human pro
16	44	50.6	1289	4	Aam79447 Human pro
17	44	50.6	1356	7	Adc06830 Human pro
18	44	50.6	1593	7	Adc06829 Human pro
19	44	50.6	1891	4	Abg04331 Novel hum
20	43	49.4	29	3	Abp40987 Human ret
21	43	49.4	126	6	Abp80192 N. gonorr
22	43	49.4	360	3	Abp40989 Human ret
23	43	49.4	420	3	Abp40990 Human ret
24	43	49.4	433	6	Abp78342 N. gonorr
25	43	49.4	459	3	Aay75225 Neisseria

26	43	49.4	459	3	AAY75223	Aay75223 Neisseria
27	43	49.4	465	6	ABU63641	Abu63641 BsmAI res
28	43	49.4	466	6	ABU32026	Abu32026 Protein e
29	43	49.4	484	7	ABO63378	AbO63378 Klebsiell
30	43	49.4	530	6	ABR43482	AbR43482 Bacillus
31	42	48.3	280	6	ABM68845	Abm68845 Photorhab
32	42	48.3	652	4	ABB67974	Abb67974 Drosoephil
33	42	48.3	951	2	AAV34536	Aay34536 Porphorym
34	42	48.3	953	2	AAV34403	Aay34403 Porphorym
35	41	47.1	111	4	ABG10728	Abg10728 Novel hum
36	41	47.1	111	4	ABG25967	Abg25967 Novel hum
37	41	47.1	144	7	ADF05142	Adf05142 Bacterial
38	41	47.1	227	7	ABO76461	AbO76461 Pseudomon
39	41	47.1	497	7	ADI27244	Adi27244 B. napus
40	41	47.1	544	8	ADG98169	Adg98169 BsaI res
41	41	47.1	588	4	AAU27909	Aau27909 Human con
42	41	47.1	651	4	ABG08359	Abg08359 Novel hum
43	41	47.1	692	8	ADJ49464	Adj49464 Oil-aesoc
44	41	47.1	817	8	ADJ49477	Adj49477 Oil-aesoc
45	40	46.0	87	4	ABG08529	Abg08529 Novel hum
46	40	46.0	223	5	AAO17425	Aao17425 A. thalian
47	40	46.0	223	7	ADH89344	Adh89344 A. thalian
48	40	46.0	249	3	ABG06462	Abg06462 Arabidops
49	40	46.0	266	3	ABG06461	Abg06461 Arabidops
50	40	46.0	311	3	ABG06460	Abg06460 Arabidops
51	40	46.0	334	2	AAV32076	Aay32076 Bacillus
52	40	46.0	346	6	ADB10514	Adb10514 Alloiooc
53	40	46.0	351	6	ADB10512	Adb10512 Alloiooc
54	40	46.0	618	7	ADB31933	Adb31933 Plant (A.
55	40	46.0	619	5	AAU93033	Aau93033 Arabidops
56	40	46.0	619	7	ADD30861	AdD30861 Plant yie
57	40	46.0	619	8	ADI41701	Adi41701 Plant tra
58	40	46.0	619	8	ADO02153	AdO02153 Thalecres
59	40	46.0	640	4	ABG10625	Abg10625 Novel hum
60	40	46.0	920	4	ABE09955	Abe09955 Methylo
61	40	46.0	920	5	ABP61576	Abp61576 High grow
62	39	44.8	87	5	ABP08056	Abp08056 Human ORF
63	39	44.8	198	4	AAU69448	Aau69448 Human pur
64	39	44.8	198	5	ABB97749	Abb97749 Human sec
65	39	44.8	224	6	ABU21012	Abu21012 Protein e
66	39	44.8	229	2	AAW05051	Aaw05051 Human thr
67	39	44.8	263	4	ABG04508	Abg04508 Novel hum
68	39	44.8	263	4	ABG17554	Abg17554 Novel hum
69	39	44.8	263	4	AAU32821	Aau32821 Novel hum
70	39	44.8	282	2	AAU97701	Aau97701 Human MPL
71	39	44.8	282	2	AAW05045	Aaw05045 Human thr
72	39	44.8	286	2	AAW08027	Aaw08027 Human MGD
73	39	44.8	286	2	AAW05044	Aaw05044 Human pla
74	39	44.8	286	2	AAW05044	Aaw05044 Human thr
75	39	44.8	286	2	AAW05044	Aaw05044 Human MGD
76	39	44.8	286	2	AAW05044	Aaw05044 Human MGD
77	39	44.8	286	3	AAV51118	Aay51118 Human pla
78	39	44.8	286	3	AAV51117	Aay51117 Human pla
79	39	44.8	313	8	ADS25343	AdS25343 Bacterial
80	39	44.8	313	8	ADS22766	AdS22766 Bacterial
81	39	44.8	313	8	ADS25853	AdS25853 Bacterial
82	39	44.8	313	8	ADS25536	AdS25536 Bacterial
83	39	44.8	352	4	AAU47738	Aau47738 Propionib
84	39	44.8	352	6	ABM44257	Abm44257 Propionib
85	39	44.8	356	2	AAV34471	Aay34471 Porphorym
86	39	44.8	433	4	ABG98942	Abg98942 E. coli g
87	39	44.8	433	6	ABU14777	Abu14777 Protein e
88	39	44.8	458	5	ABP65946	Abp65946 Bifidobac
89	39	44.8	491	2	AAV34347	Aay34347 Porphorym
90	39	44.8	671	4	ABG17780	Abg17780 Novel hum
91	39	44.8	696	4	ABB60920	Abb60920 Drosoephil
92	39	44.8	743	4	ABG23278	Abg23278 Novel hum
93	39	44.8	743	4	ABG14502	Abg14502 Novel hum
94	39	44.8	797	4	ABG29533	Abg29533 Novel hum
95	39	44.8	924	8	ADP99017	Adp99017 C. albica
96	39	44.8	978	6	ABU26580	Abu26580 Protein e
97	39	44.8	1420	8	ADR86081	Adr86081 Aspergill
98	39	44.8	1647	4	ABG10750	Abg10750 Novel hum

99 39 44.8 2104 2 AAW57773 Aaw57773 Musca dom
100 39 44.8 2104 2 AAW89578 Aaw89578 Calcium p

ALIGNMENTS

RESULT 1
AAAY30352
ID AAY30352 standard; peptide; 15 AA.

XX AC AAY30352;
XX DT 09-NOV-1999 (first entry)
XX DE Epitope derived from pneumococcal surface adhesion A protein.
XX KW Pneumococcal surface adhesion A protein; PsaA; monoclonal antibody;
XX KW vaccine; Streptococcus pneumoniae infection.

XX OS Streptococcus pneumoniae.

XX PN WO9945121-A1.

XX PD 10-SEP-1999.

XX PF 26-FEB-1999; 99WO-US004326.

XX PR 02-MAR-1998; 98US-0076565P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Carlone GW, Ades EW, Sampson JS, Tharpe JA, Zeiler JL;

XX PI Westerink MAJ;

XX DR WPI; 1999-540849/45.

XX PT New peptides corresponding to Streptococcus pneumoniae PsaA, used for
XX PT treating or preventing Streptococcus pneumoniae infection in a subject.

XX PS Claim 6; Page 43; 58pp; English.

XX CC AAY30351-54 represent immunogenic peptides which are derived from a
XX CC pneumococcal surface adhesion A protein (PsaA). The specification
XX CC describes monoclonal antibodies which bind epitopes of the PsaA protein
XX CC (e.g present sequence). The peptides can be used in vaccines to prevent
XX CC Streptococcus pneumoniae infections. The antibodies of the invention can
XX CC also be used to detect S. pneumoniae in a sample or individual

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 87; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFWRLL 15
DB 1 RSYQHDRLRAYGFWRLL 15

RESULT 2
AAE19240
ID AAE19240 standard; peptide; 15 AA.

XX AC AAE19240;

XX DT 21-MAY-2002 (first entry)

XX DE Streptococcus pneumoniae PsaA immunogenic peptide #2.

XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.

OS Streptococcus pneumoniae.
XX PN WO200204497-A2.
XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GW;
XX DR WPI; 2002-195762/25.
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX PT fragment.

XX PS Claim 2; Page 56; 86pp; English.

XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is Streptococcus pneumoniae PsaA immunogenic peptide

XX SQ Sequence 15 AA;

Query Match 100.0%; Score 87; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFWRLL 15
DB 1 RSYQHDRLRAYGFWRLL 15

RESULT 3

AAE19245

ID AAE19245 standard; protein; 66 AA.

XX AC AAE19245;

XX DT 21-MAY-2002 (first entry)

XX DE Bi-peptide heterogenous multiple antigenic peptide (MAP) #1.

XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.

XX OS Unidentified.

XX FH Key Location/Qualifiers
XX FT Peptide 1..15 /note= "PsaA immunogenic peptide #1"
XX FT Peptide 16..30 /note= "PsaA immunogenic peptide #2"
XX FT Peptide 31..45 /note= "PsaA immunogenic peptide #1"
XX FT Peptide 46..60 /note= "PsaA immunogenic peptide #2"
XX FT Modified-site 62 /label= Nle
XX FT Modified-site 64 /label= Nle
XX FT Modified-site 66 /label= Nle

XX PN WO200204497-A2.

```

XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX PT fragment.
XX PS Example 1b; Fig 1A; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is a bi-peptide heterogenous MAP used to provide
XX CC protection against Streptococcus pneumoniae challenge
XX SQ Sequence 66 AA;

Query Match 100.0%; Score 87; DB 5; Length 66;
Best Local Similarity 100.0%; Pred. No. 7.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFWR 15
Db 16 RSYQHDLRAYGFWR 30

RESULT 4
AAE19244
ID AAE19244 standard; peptide; 15 AA.
AC AAE19244;
XX DT 21-MAY-2002 (first entry)
XX DE Streptococcus pneumoniae PsaA immunogenic peptide #6.
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX OS Streptococcus pneumoniae.
XX PN WO200204497-A2.
XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX CC New multiple antigenic peptide for immunizing against streptococcal
XX CC infections, binds to monoclonal antibody obtained in response to
XX CC immunizing an animal with pneumococcal surface adhesion protein A or its
XX CC fragment.

Query Match 100.0%; Score 87; DB 5; Length 66;
Best Local Similarity 100.0%; Pred. No. 7.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFWR 15
Db 16 RSYQHDLRAYGFWR 30

RESULT 5
AAE19247
ID AAE19247 standard; protein; 47 AA.
XX AC AAE19247;
XX DT 21-MAY-2002 (first entry)
XX DE Tri-peptide heterogenous multiple antigenic peptide (MAP).
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT Peptide 1..15 /note= "PsaA immunogenic peptide #1"
XX FT Peptide 16..29 /note= "PsaA immunogenic peptide #6"
XX FT Peptide 30..43 /note= "PsaA immunogenic peptide #5"
XX FT Modified-site 45 /label= Nle
XX FT Modified-site 47 /label= Nle
XX PN WO200204497-A2.
XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX PT fragment.
XX PS Example 1b; Fig 1C; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is a bi-peptide heterogenous MAP used to provide
XX CC protection against Streptococcus pneumoniae challenge
XX SQ Sequence 15 AA;

Query Match 94.3%; Score 82; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SYQHDLRAYGFWR 15
Db 1 SYQHDLRAYGFWR 14

RESULT 5
AAE19247
ID AAE19247 standard; protein; 47 AA.
XX AC AAE19247;
XX DT 21-MAY-2002 (first entry)
XX DE Tri-peptide heterogenous multiple antigenic peptide (MAP).
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX KW pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT Peptide 1..15 /note= "PsaA immunogenic peptide #1"
XX FT Peptide 16..29 /note= "PsaA immunogenic peptide #6"
XX FT Peptide 30..43 /note= "PsaA immunogenic peptide #5"
XX FT Modified-site 45 /label= Nle
XX FT Modified-site 47 /label= Nle
XX PN WO200204497-A2.
XX PD 17-JAN-2002.
XX PF 10-JUL-2001; 2001WO-US021626.
XX PR 10-JUL-2000; 2000US-00613092.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX DR WPI; 2002-195762/25.
XX PT New multiple antigenic peptide for immunizing against streptococcal
XX PT infections, binds to monoclonal antibody obtained in response to
XX PT immunizing an animal with pneumococcal surface adhesion protein A or its
XX PT fragment.
XX PS Example 1b; Fig 1C; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
XX CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX CC obtained in response to immunising an animal with pneumococcal surface
XX CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX CC protective immunity against S. pneumoniae infection in a subject. The
XX CC present sequence is a bi-peptide heterogenous MAP used to provide
XX CC protection against Streptococcus pneumoniae challenge

```

CC present sequence is a tri-peptide heterogeneous MAP used to provide
 CC protection against Streptococcus pneumoniae challenge

XX
 SQ Sequence 47 AA;
 Query Match 94.3%; Score 82; DB 5; Length 47;
 Best Local Similarity 100.0%; Pred. No. 3.7e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGFWRLL 15
 |||||
 Db 16 SYQHDLRAYGFWRLL 29

RESULT 6
 ABG19785
 ID ABG19785 standard; protein; 554 AA.
 XX
 AC ABG19785;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19776.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS83972.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 50144; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 554 AA;
 Query Match 54.0%; Score 47; DB 4; Length 554;
 Best Local Similarity 53.8%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWRLL 15
 |::|
 Db 7 YKSLHAYGAWAI 19

RESULT 7
 ABG19731
 ID ABG19731 standard; protein; 554 AA.
 XX
 AC ABG19731;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19722.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS83918.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 50090; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Query Match 54.0%; Score 47; DB 4; Length 554;
 Best Local Similarity 53.8%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRLL 15
 | : | | | | | :
 DB 7 YKHSLSLAYGAWAI 19

RESULT 8
 ABG03083
 ID ABG03083 standard; protein; 1007 AA.
 XX
 AC ABG03083;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #3074.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS67270.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 33442; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1007 AA;

Query Match 54.0%; Score 47; DB 4; Length 1007;

Best Local Similarity 53.8%; Pred. No. 58;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRLL 15
 | : | | | | | :
 DB 7 YKHSLSLAYGAWAI 19

RESULT 9
 ABP79868
 ID ABP79868 standard; protein; 741 AA.
 XX
 AC ABP79868;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae amino acid sequence SEQ ID 6266.
 XX
 KW Antibacterial; infection; vaccine; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 FN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB002069.
 XX
 PR 12-FEB-2001; 2001GB-00003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizza M, Massignani V, Monaci E;
 XX
 DR WPI; 2003-058415/05.
 DR N-PSDB; ABZ40838.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 PS Disclosure; Page 645; 815pp; English.
 XX
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 741 AA;

Query Match 51.7%; Score 45; DB 6; Length 741;
 Best Local Similarity 58.3%; Pred. No. 91;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 QHDLRAYGFWRLL 15
 | : | | | | | :
 DB 448 QHDEAGGIWRM 459

RESULT 10
 ADF03901
 ID ADF03901 standard; protein; 903 AA.
 XX
 AC ADF03901;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Bacterial polypeptide #14.
 XX
 KW Proteus mirabilis infection; bacterial infection; antibacterial;

KW immunostimulant.
 XX
 OS Proteus mirabilis.
 XX
 XX US6605709-B1.
 XX
 XX 12-AUG-2003.
 PD
 XX
 XX 05-APR-2000; 2000US-00543681.
 PF
 XX
 XX 09-APR-1999; 99US-0128706P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX
 XX Breton GL;
 FI
 XX WPI; 2003-895291/82.
 XX
 DR N-PSDB; ADE99728.
 DR
 XX
 XX New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX
 XX Disclosure; SEQ ID NO 4186; 870pp; English.
 PS
 XX The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polypeptide of the invention.
 XX
 SQ Sequence 903 AA;
 Query Match 51.7%; Score 45; DB 7; Length 903;
 Best Local Similarity 53.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 SYQHDRLRAYGFWR 14
 Db 110 NYRHRLGYGFWR 122
 :||| |
 RESULT 11
 ABP02252
 ID ABP02252 standard; protein; 118 AA.
 XX
 XX ABP02252;
 AC
 XX 24-JUN-2002 (first entry)
 DT
 XX Human ORFX protein sequence SEQ ID NO:4486.
 DE
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 XX WO200192523-A2.
 PN
 XX

PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US010836.
 XX
 XX 30-MAY-2000; 2000US-0206132P.
 PR 29-AUG-2000; 2000US-0228716P.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shinkets RA, Leach MD;
 FI WPI; 2002-106308/14.
 XX
 DR N-PSDB; ABN18004.
 DR
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX
 XX Disclosure; SEQ ID NO 4486; 1037pp; English.
 PS
 XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative diseases, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 118 AA;
 Query Match 50.6%; Score 44; DB 5; Length 118;
 Best Local Similarity 40.0%; Pred. No. 20;
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 RSYQHDRLRAYGFWRL 15
 Db 2 RTFDYDVDARGFWQI 16
 :||| |
 RESULT 12
 ABG26189
 ID ABG26189 standard; protein; 267 AA.
 XX
 XX ABG26189;
 AC
 XX 18-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #26180.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX

PD	11-OCT-2001.	
XX	30-MAR-2001; 2001WO-US008631.	
XX	31-MAR-2000; 2000US-00540217.	
XX	23-AUG-2000; 2000US-00649167.	
XX	(HYSE-) HYSEQ INC.	
XX	Dmanac RT, Liu C, Tang YT;	
XX	WPI; 2001-639362/73.	
XX	N-PSDB; AAS90376.	
XX	New isolated polynucleotide and encoded polypeptides, useful in	
XX	diagnostics, forensics, gene mapping, identification of mutations	
XX	responsible for genetic disorders or other traits and to assess	
XX	biodiversity.	
XX	Claim 20; SEQ ID NO 56548; 103pp; English.	
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)	
XX	sequences. (I) is useful as hybridisation probes, polymerase chain	
XX	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	
XX	and in recombinant production of (II). The polynucleotides are also used	
XX	in diagnostics as expressed sequence tags for identifying expressed	
XX	genes. (II) is useful in gene therapy techniques to restore normal	
XX	activity of (II) or to treat disease states involving (II). (II) is	
XX	useful for generating antibodies against it, detecting or quantitating a	
XX	polypeptide in tissue, as molecular weight markers and as a food	
XX	supplement. (II) and its binding partners are useful in medical imaging	
XX	of sites expressing (II). (I) and (II) are useful for treating disorders	
XX	involving aberrant protein expression or biological activities in	
XX	polypeptide and polynucleotide sequences have applications in	
XX	diagnostics, forensics, gene mapping, identification of mutations	
XX	responsible for genetic disorders or other traits to assess biodiversity	
XX	and to produce other types of data and products dependent on DNA and	
XX	amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic	
XX	amino acid sequences of the invention. Note: The sequence data for this	
XX	patent did not appear in the printed specification, but was obtained in	
XX	electronic format directly from WIPO at	
XX	ftp.wipo.int/pub/published_pct_sequences	
XX	Sequence 267 AA;	
SQ	Query Match 50.6%; Score 44; DB 4; Length 267;	
	Best Local Similarity 53.8%; Pred. No. 47;	
	Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	
QY	1 RSYQHDLRAYGFW 13	
DB	120 RAYRADLRLEGYW 132	
RESULT 13		
ID	AAE25282	
XX	AAE25282 standard; protein; 1096 AA.	
XX	AAE25282;	
XX	30-OCT-2002 (first entry)	
XX	Human nucleic acid-associated protein (NAAP-1).	
XX	Human; nucleic acid-associated protein; NAAP-1; neurological disorder;	
XX	arteriosclerosis; cancer; cell proliferative disorder; arteriosclerosis;	
XX	lymphoma; epilepsy; Alzheimer's disease; developmental; anticonvulsant;	
XX	autoimmune disorder; AIDS; allergy; anaemia; stroke; malaria; leishmania;	
XX	gene therapy; nontropic; neuroprotective; cerebroprotective; virucide;	
XX	immunosuppressive; protozoacide; antimicrobial.	
XX	Homo sapiens.	
OS		

PH	Key	Location/Qualifiers
FT	Domain	478..503
FT		/note= "Transmembrane domain"
FT	Domain	621..649
FT		/note= "Transmembrane domain"
XX	WO200250279-A2.	
XX	27-JUN-2002.	
XX	19-DEC-2001; 2001WO-US050256.	
XX	21-DEC-2000; 2000US-0257714P.	
XX	05-JAN-2001; 2001US-0260081P.	
XX	15-JAN-2001; 2001US-0262302P.	
XX	23-JAN-2001; 2001US-0263923P.	
XX	02-FEB-2001; 2001US-0266088P.	
XX	29-OCT-2001; 2001US-0348442P.	
XX	(INCY-) INCYTE GENOMICS INC.	
XX	Baughin MR, Lu Y, Arvizu C, Ramkumar J, Yao MG, Policky JL;	
XX	Walia NK, Tribouley KM, Yue H, Batra S, Ding L, Lal PG;	
XX	Borowsky ML, Lu DAM, Gandhi AR, Griffin JA, Xu Y, Azimzai Y;	
XX	Gietzen KJ, Tang YT, Warren BA, Mason PM, Burford N, Hafalia AJA;	
XX	Lee EA, Yang J, Gorvad AE, Emerling BM, Marquis JP, Lee SY;	
XX	Swarnakar A, Reddy R;	
XX	WPI; 2002-519887/55.	
XX	N-PSDB; AAD41191.	
XX	Nucleic acid associated proteins and nucleic acids for diagnosing,	
XX	treating and preventing cell proliferative (e.g. cancers), neurological	
XX	(e.g. epilepsy or stroke) or autoimmune disorders (e.g. AIDS).	
XX	Claim 56; Page 151-154; 193pp; English.	
XX	The invention relates to nucleic acid-associated proteins (NAAP) and	
XX	nucleic acids. The nucleic acid and amino acid sequences are useful for	
XX	diagnosing, treating and preventing cell proliferative e.g.	
XX	arteriosclerosis, atherosclerosis, lymphoma or cancers), neurological	
XX	(e.g. epilepsy, Alzheimer's disease or stroke), developmental, and	
XX	autoimmune disorders (e.g. AIDS, allergies, or anaemia) or infections	
XX	(e.g. malaria, or leishmania), as well as in assessing the effects of	
XX	exogenous compound on the expression of nucleic acid and amino acid	
XX	sequences of nucleic acid-associated proteins. The invention is useful in	
XX	gene therapy. The present sequence is human NAAP-1	
XX	Sequence 1096 AA;	
SQ	Query Match 50.6%; Score 44; DB 5; Length 1096;	
	Best Local Similarity 58.3%; Pred. No. 28+02;	
	Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	
QY	3 YQHDRLRAYGFW 14	
DB	20 YQHYRLYGQWK 31	
RESULT 14		
ID	AAB08517	
XX	AAB08517 standard; protein; 1114 AA.	
XX	AAB08517;	
XX	20-DEC-2000 (first entry)	
XX	Amino acid sequence of a beta-tubulin antigen.	
XX	Beta-tubulin antigen; inner ear protein; Meniere's disease; autoantibody;	
XX	chronic ear disease; autoimmune disease.	
XX	Homo sapiens.	
OS		

(HYSE-) HYSEQ INC.
 Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 WPI; 2001-476283/51.
 N-PSDB; AAKS2580.
 Nucleic acids encoding polypeptides with cytokine-like activities, useful
 in diagnosis and gene therapy.
 Claim 20; Page 250-251; 6221pp; English.
 The invention relates to polynucleotides (AAK51456-AAK53435) and the
 encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 cytokine, cell proliferation or cell differentiation or which may induce
 production of other cytokines in other cell populations. The
 polynucleotides and polypeptides are useful in gene therapy, vaccines or
 peptide therapy. The polypeptides have various cytokine-like activities,
 e.g. stem cell growth factor activity, haematopoiesis regulating
 activity, tissue growth factor activity, immunomodulatory activity and
 activin/inhibin activity and may be useful in the diagnosis and/or
 treatment of cancer, leukaemia, nervous system disorders, arthritis and
 inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 sequence listing were missing at the time of publication

Sequence 1289 AA;
 Query Match 50.6%; Score 44; DB 4; Length 1289;
 Best Local Similarity 58.3%; Pred. No. 2.4e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFWR 14
 |||||
 Db 205 YQHRVRLYGQWK 216
 |||||

RESULT 17
 ADC06830
 ID ADC06830 standard; protein; 1356 AA.
 AC ADC06830;
 DT 18-DEC-2003 (first entry)
 DE Human prostate cancer-related protein AF441770.
 XX cytostatic; prostate cancer; breast; gene therapy; transgenic; human.
 KW Homo sapiens.
 OS
 XX WO2003064599-A2.
 PN
 XX 07-AUG-2003.
 PD
 XX
 XX 24-JAN-2003; 2003WO-US001943.
 PF
 XX 25-JAN-2002; 2002US-00054935.
 PR 14-FEB-2002; 2002US-0356130P.
 PR 22-MAR-2002; 2002US-00102946.
 PR 08-APR-2002; 2002US-00117229.
 PR 14-MAY-2002; 2002US-00144198.
 PR 19-JUL-2002; 2002US-00197824.
 XX
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX Sun Z, Li X, Jay G, Kovacs KP, Fan W, Shu Y;
 PI WPI; 2003-679495/64.
 XX New isolated polynucleotide related to cancer genes, useful for

PT detecting, diagnosing, staging, monitoring, prognosticating, preventing
 or treating cancers, e.g. breast and prostate cancers.
 PS Disclosure; Fig 15; 128pp; English.
 XX The invention relates to a novel isolated polynucleotide comprising a
 differentially-regulated mammalian cancer gene. The polynucleotides of
 the invention demonstrate cytostatic activity and are differentially
 expressed in prostate cancer. The polynucleotide, polypeptides and
 methods of the invention may be useful for detecting, diagnosing,
 staging, monitoring, prognosticating, preventing or treating cancers,
 particularly breast and prostate cancers. Furthermore, the invention may
 be utilised during gene therapy procedures or in the production of
 transgenic animals. The current sequence is that of the prostate cancer-
 related protein of the invention.

Sequence 1356 AA;
 Query Match 50.6%; Score 44; DB 7; Length 1356;
 Best Local Similarity 58.3%; Pred. No. 2.5e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFWR 14
 |||||
 Db 280 YQHRVRLYGQWK 291
 |||||

RESULT 18
 ADC06829
 ID ADC06829 standard; protein; 1593 AA.
 AC ADC06829;
 DT 18-DEC-2003 (first entry)
 DE Human prostate cancer-related protein PCP0557.
 XX cytostatic; prostate cancer; breast; gene therapy; transgenic; human;
 KW chromosome Xq25-q26.3.
 OS Homo sapiens.
 XX WO2003064599-A2.
 PN
 XX 07-AUG-2003.
 PD
 XX 24-JAN-2003; 2003WO-US001943.
 PF
 XX 25-JAN-2002; 2002US-00054935.
 PR 14-FEB-2002; 2002US-0356130P.
 PR 22-MAR-2002; 2002US-00102946.
 PR 08-APR-2002; 2002US-00117229.
 PR 14-MAY-2002; 2002US-00144198.
 PR 19-JUL-2002; 2002US-00197824.
 XX
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX Sun Z, Li X, Jay G, Kovacs KP, Fan W, Shu Y;
 PI WPI; 2003-679495/64.
 XX New isolated polynucleotide related to cancer genes, useful for
 detecting, diagnosing, staging, monitoring, prognosticating, preventing
 or treating cancers, e.g. breast and prostate cancers.
 PS Claim 5; Fig 15; 128pp; English.
 XX The invention relates to a novel isolated polynucleotide comprising a
 differentially-regulated mammalian cancer gene. The polynucleotides of
 the invention demonstrate cytostatic activity and are differentially
 expressed in prostate cancer. The polynucleotide, polypeptides and
 methods of the invention may be useful for detecting, diagnosing,
 staging, monitoring, prognosticating, preventing or treating cancers,

RESULT 21
ID ABP80192 standard; protein; 126 AA.
AC ABP80192;
DT 07-MAR-2003 (first entry)
DE N. gonorrhoeae amino acid sequence SEQ ID 6914.
KW Antibacterial; infection; vaccine; gene therapy.
OS Neisseria gonorrhoeae.
XX WO200279243-A2.
XX 10-OCT-2002.
XX 12-FEB-2002; 2002WO-IB002069.
XX 12-FEB-2001; 2001GB-00003424.
XX (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Masignani V, Monaci E;
XX WPI; 2003-058415/05.
XX N-PSDB; ABZ41162.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
XX Disclosure; Page 689; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX molecules of the invention
XX Sequence 126 AA;
SQ

Query Match 49.4%; Score 43; DB 6; Length 126;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 6 DLRAYGFWRLL 15
DB 94 DLGAYGFWSM 103

RESULT 22
ID ABP40989 standard; protein; 360 AA.
AC ABP40989;
XX 01-AUG-2002 (first entry)
DE Human retroviral gag protein.
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis.
XX Human endogenous retrovirus.
XX WO9967395-A1.
XX 29-DEC-1999.
XX

Query Match 49.4%; Score 43; DB 3; Length 360;
Best Local Similarity 53.8%; Pred. No. 93;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 3 YQHDRLRAYGFWRLL 15
DB 336 FSHDPQASGLWRLL 348

RESULT 23
ID ABP40990 standard; protein; 420 AA.
AC ABP40990;
XX 01-AUG-2002 (first entry)
DE Human retroviral gag protein.
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis.
XX Human endogenous retrovirus.
XX WO9967395-A1.
XX 29-DEC-1999.
XX 23-JUN-1999; 99WO-FR001513.
XX 23-JUN-1998; 98FR-00007920.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX for diagnosis, treatment and prevention of autoimmune and neurological
XX diseases.
XX

XX 23-JUN-1999; 99WO-FR001513.
XX 23-JUN-1998; 98FR-00007920.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX for diagnosis, treatment and prevention of autoimmune and neurological
XX diseases.
XX Claim 22; Page 154-155; 225pp; French.
XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
XX sequences can be used in immunogenic or vaccinating compositions, for
XX protection against autoimmune diseases, particularly multiple sclerosis.
XX The peptides may also be used (by sequence comparison) to detect/identify
XX endogenous retroviruses that are abnormally expressed in cancer.
XX neuropathologies or other autoimmune diseases. The present sequence was
XX used to illustrate the invention
XX Sequence 360 AA;
SQ

Query Match 49.4%; Score 43; DB 3; Length 360;
Best Local Similarity 53.8%; Pred. No. 93;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 3 YQHDRLRAYGFWRLL 15
DB 336 FSHDPQASGLWRLL 348

RESULT 23
ID ABP40990 standard; protein; 420 AA.
AC ABP40990;
XX 01-AUG-2002 (first entry)
DE Human retroviral gag protein.
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis.
XX Human endogenous retrovirus.
XX WO9967395-A1.
XX 29-DEC-1999.
XX 23-JUN-1999; 99WO-FR001513.
XX 23-JUN-1998; 98FR-00007920.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX for diagnosis, treatment and prevention of autoimmune and neurological
XX diseases.
XX

PT diseases.

XX Claim 22; Fig 6; 225pp; French.

XX The present invention relates to new nucleic acid sequences of human

CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.

CC Regulatory elements associated with HERV-7q may alter expression of other

CC genes (even remote genes) on the same chromosome, inducing immunological

CC and/or neurological changes (which may be pathological or protective/

CC curative). HERV-7q peptides can be used to improve efficiency of the

CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding

CC sequences can be used in immunogenic or vaccinating compositions for

CC protection against autoimmune diseases, particularly multiple sclerosis.

CC The peptides may also be used (by sequence comparison) to detect/identify

CC endogenous retroviruses that are abnormally expressed in cancer,

CC neuropathologies or other autoimmune diseases. The present sequence was

CC used to illustrate the invention

XX Sequence 420 AA;

SQ

Query Match 49.4%; Score 43; DB 3; Length 420;

Best Local Similarity 53.8%; Pred. No. 1.1e+02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRL 15

DB 337 FSHDPQASGLWRL 349

RESULT 24

ABP78342

ID ABP78342 standard; protein; 433 AA.

AC ABP78342;

XX 07-MAR-2003 (first entry)

XX N. gonorrhoeae amino acid sequence SEQ ID 3214.

XX Antibacterial; infection; vaccine; gene therapy.

XX Neisseria gonorrhoeae.

XX WO200279243-A2.

XX 10-OCT-2002.

XX 12-FEB-2002; 2002WO-IB002069.

XX 12-FEB-2001; 2001GB-00003424.

XX (CHIR-) CHIRON SPA.

XX Fontana MR, Pizza M, Maignani V, Monaci E;

XX WPI; 2003-058415/05.

XX N-PSDB; ABZ39312.

XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a

XX medicament for treating or preventing N. gonorrhoeae infection.

XX Disclosure; Page 424; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.

XX Also disclosed are the nucleic acid molecules encoding the proteins and

XX antibodies that specifically bind to the proteins. The composition

XX comprising the protein, nucleic acid or antibody is useful for the

XX manufacture of a medicament for treating or preventing N. gonorrhoeae

XX infection, this may be in the form of a vaccine or gene therapy.

XX Sequences given in records ABP76736-ABP81046 represent nucleic acid

XX molecules of the invention

XX Sequence 433 AA;

Query Match 49.4%; Score 43; DB 3; Length 459;

SQ

Query Match 49.4%; Score 43; DB 6; Length 433;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 13

DB 390 AYRFDWGIYGF 401

RESULT 25

AAV75225

ID AAV75225 standard; protein; 459 AA.

XX AAV75225;

XX 21-MAR-2000 (first entry)

XX Neisseria meningitidis ORF 607 protein sequence SEQ ID NO:1924.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

XX antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;

XX antibacterial; gene therapy.

XX Neisseria meningitidis.

XX WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US009346.

XX 01-MAY-1998; 98US-0083758P.

XX 31-JUL-1998; 98US-0094869P.

XX 02-SEP-1998; 98US-0098994P.

XX 02-SEP-1998; 98US-0099062P.

XX 09-OCT-1998; 98US-0103749P.

XX 09-OCT-1998; 98US-0103794P.

XX 09-OCT-1998; 98US-0103796P.

XX 25-FEB-1999; 99US-0121528P.

XX (CHIR) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;

XX Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

XX Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

XX N-PSDB; AAZ53987.

XX Novel Neisserial polypeptides predicted to be useful antigens for

XX vaccines and diagnostics.

XX Claim 2; Page 969; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ54253 to AAZ54941

XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides

XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent

XX PCR primers used in the exemplification of the present invention. The

XX polypeptides, the polynucleotides, antibodies and compositions of the

XX invention can be used as vaccines, as diagnostic reagents, and as

XX immunogenic compositions. The polypeptides can be used in the manufacture

XX of medicaments for treating or preventing infection due to Neisserial

XX bacteria (e.g. meningitis and septicemia), to detect the presence of

XX Neisseria bacteria, or to raise antibodies. They may also be used to

XX screen for agonists or antagonists, which may themselves have use as

XX antibacterial agents. The polynucleotides of the invention may also be

XX used in gene therapy protocols

XX Sequence 459 AA;

Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGF 13
Db 416 AYRFDGMIYGF 427

RESULT 26
AAV75223
ID AAV75223 standard; protein; 459 AA.

XX AC AAV75223;

XX 12-SEP-2003 (revised)
DT 21-MAR-2000 (first entry)

XX Neisseria gonorrhoeae ORF 607 protein sequence SEQ ID NO:1920.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US009346.

XX 01-MAY-1998; 98US-0083758P.

XX 31-JUL-1998; 98US-0094869P.

XX 02-SEP-1998; 98US-0098994P.

XX 02-SEP-1998; 98US-0099062P.

XX 09-OCT-1998; 98US-0103749P.

XX 09-OCT-1998; 98US-0103794P.

XX 09-OCT-1998; 98US-0103796P.

XX 25-FEB-1999; 99US-0121528P.

XX (GENO-) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzi M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;

XX WPI; 2000-062150/05.

XX N-PSDB; AAZ53985.

XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.

XX Claim 2; Page 967; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAY75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to Neisseria
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
CC field)

XX Sequence 459 AA;

Query Match 49.4%; Score 43; DB 3; Length 459;

Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGF 13
Db 416 AYRFDGMIYGF 427

RESULT 27
ABU63641
ID ABU63641 standard; protein; 465 AA.

XX AC ABU63641;

XX 23-OCT-2003 (revised)
DT 30-SEP-2003 (first entry)

XX BsmAI restriction endonuclease, BsmAIR.

XX Restriction endonuclease; enzyme; AsiSI; DNA cleaving; methylation;
KW molecular cloning; BsmAIR.

XX OS Geobacillus stearothermophilus.

XX US2003104388-A1.

XX 05-JUN-2003.

XX 20-SEP-2001; 2001US-00957005.

XX 20-SEP-2001; 2001US-00957005.

XX (NEWE) NEW ENGLAND BIOLABS INC.

XX Zhu Z, Zhou J, Xu S;

XX WPI; 2003-567952/53.

XX N-PSDB; ACD28244.

XX New DNA coding for the AsiSI restriction endonuclease or AsiSI methylase,
PT useful for cleaving DNA molecules into small fragments for molecular
PT cloning and gene characterization.

XX Disclosure; Fig 3; 27pp; English.

XX The invention relates to a new isolated DNA coding for the AsiSI
CC restriction endonuclease or AsiSI methylase. The DNA is useful for
CC cleaving DNA molecules into small fragments for molecular cloning and
CC gene characterisation. The DNA is also useful for modifying DNA molecules
CC via methylation. The present sequence represents the amino acid sequence
CC of the BsmAI restriction endonuclease, BsmAIR. (Updated on 23-OCT-2003 to
CC standardise OS field)

XX Sequence 465 AA;

Query Match 49.4%; Score 43; DB 6; Length 465;

Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 13
Db 214 KSYTKDRRAYEY 226

RESULT 28
ABU32026
ID ABU32026 standard; protein; 466 AA.

XX AC ABU32026;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #17553.

```
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Klebsiella pneumoniae.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA35896.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 59950; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 466 AA;
XX
XX Query Match 49.4%; Score 43; DB 6; Length 466;
XX Best Local Similarity 87.5%; Pred. No. 1.2e+02;
XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 8 RAYGFWRLL 15
XX Db 430 RRYGFWRLL 437
XX
XX RESULT 29
XX ABO63378
XX ID ABO63378 standard; protein; 484 AA.
XX AC ABO63378;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polypeptide seqid 9895.
XX KW Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX OS Klebsiella pneumoniae.
XX PN US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX WPI; 2003-895346/82.
XX N-PSDB; ACH96929.
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 9895; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention
XX Sequence 484 AA;
XX
XX Query Match 49.4%; Score 43; DB 7; Length 484;
XX Best Local Similarity 87.5%; Pred. No. 1.3e+02;
XX Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 8 RAYGFWRLL 15
XX Db 448 RRYGFWRLL 455
XX
XX RESULT 30
XX ABR43482
XX ID ABR43482 standard; protein; 530 AA.
XX AC ABR43482;
XX DT 23-OCT-2003 (revised)
XX DT 16-JUL-2003 (first entry)
XX DE Bacillus stearothermophilus B61 endonuclease protein SEQ ID NO:13.
XX KW Bacillus stearothermophilus B61; BsmI; endonuclease; bamBIR; enzyme;
XX molecular cloning; gene characterisation.
XX OS Geobacillus stearothermophilus.
XX PN EP1298212-A2.
XX
```

PD 02-APR-2003.
 XX
 PF 27-SEP-2002; 2002EP-00256756.
 XX
 PR 28-SEP-2001; 2001US-00966997.
 XX
 PA (NEW) NEW ENGLAND BIOLABS INC.
 XX
 XX Xu S, Dore A, Hume A, Pellitier J, Zhou J;
 PI WPI; 2003-395598/38.
 XX N-PSDB; ACC69533.
 DR
 DR New DNA segment coding for the BsmBI restriction endonuclease and/or
 PT BsmBI methylase, useful for molecular cloning and gene characterization,
 PT and in producing restriction endonucleases and modification methylases.
 XX
 XX Example 1; Fig 3A-B; 35pp; English.
 PS
 XX The present sequence represents Bacillus stearothermophilus B61 BsmBI
 CC restriction endonuclease (bsmBIR). The present invention also describes
 CC BsmBI methylase (bsmBIM) isolated from B. stearothermophilus B61. BsmBIR
 CC and bsmBIM can be isolated from a DNA segment obtainable from American
 CC Type Culture Collection (ATCC) Accession Number PTA-3739. The DNA segment
 CC is useful in molecular cloning and gene characterisation, and in
 CC producing BsmBI restriction endonucleases and modification methylases.
 CC (Updated on 23-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 530 AA;
 Query Match 49.4%; Score 43; DB 6; Length 530;
 Best Local Similarity 53.8%; Pred. No. 1.4e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RSYQHDLRAYGFW 13
 | | | | | : |
 DB 226 RLYNHDRAPFWW 238
 RESULT 31
 ABM68845
 ID ABM68845 standard; protein; 280 AA.
 XX
 AC ABM68845;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens protein sequence #1942.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 XX
 OS Photorhabdus luminescens.
 XX
 XX WO200294867-A2.
 PN
 XX 28-NOV-2002.
 PD
 XX 07-FEB-2002; 2002WO-IB003040.
 PF
 XX 07-FEB-2001; 2001FR-00001659.
 PR
 XX (INSP) INST PASTEUR.
 PA
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX WPI; 2003-148459/14.
 DR
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT

PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX Claim 2; SEQ ID NO 1942; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 XX
 SQ Sequence 280 AA;
 Query Match 48.3%; Score 42; DB 6; Length 280;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 HDLRAYGFW 13
 | | | | | : |
 DB 171 HRRRSYGFW 179
 RESULT 32
 ABB67974
 ID ABB67974 standard; protein; 652 AA.
 XX
 AC ABB67974;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 30714.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 PR 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX N-PSDB; ABL12077.
 DR
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 30714; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at fp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 652 AA;
 Query Match 48.3%; Score 42; DB 4; Length 652;
 Best Local Similarity 53.8%; Pred. No. 2.5e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RSYQHDLRAYGFW 13
 ||:|:|
 Db 557 RSFQYSLTDYHF 569
 RESULT 33
 AAY34536
 ID AAY34536 standard; protein; 951 AA.
 XX AC AAY34536;
 XX DT 27-AUG-2003 (revised)
 DT 20-MAR-2003 (revised)
 DT 25-AUG-1999 (first entry)
 XX DE Porphyromonas gingivalis protein PG67.
 XX KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.
 XX OS Porphyromonas gingivalis.
 XX PN WO9929870-A1.
 XX PD 17-JUN-1999.
 XX PF 10-DEC-1998; 98WO-AU001023.
 XX PR 10-DEC-1997; 97AU-00000839.
 PR 31-DEC-1997; 97AU-00001182.
 PR 30-JAN-1998; 98AU-00001546.
 PR 10-MAR-1998; 98AU-00002264.
 PR 09-APR-1998; 98AU-00002911.
 PR 23-APR-1998; 98AU-00003128.
 PR 05-MAY-1998; 98AU-00003338.
 PR 22-MAY-1998; 98AU-00003654.
 PR 29-JUL-1998; 98AU-00004317.
 PR 30-JUL-1998; 98AU-00004963.
 PR 04-AUG-1998; 98AU-00005028.
 XX PA (CSLC-) CSL LTD.
 XX PI Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
 PI Hocking DM, Webb EA;
 XX WPI; 1999-385613/32.
 DR N-PSDB; AAX91754.
 XX PT Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
 XX Claim 1; Page 526-527; 589pp; English.
 XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation

CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation
 CC assays. Porphyromonas gingivalis is involved in periodontal disease
 CC especially gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX SQ Sequence 951 AA;
 Query Match 48.3%; Score 42; DB 2; Length 951;
 Best Local Similarity 56.2%; Pred. No. 3.7e+02;
 Matches 9; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 QY 2 SYQHDL--RAYGFWR 15
 ||:|:|
 Db 599 SYMPDFTKRRYGFWE 614
 RESULT 34
 AAY34403
 ID AAY34403 standard; protein; 953 AA.
 XX AC AAY34403;
 XX DT 27-AUG-2003 (revised)
 DT 20-MAR-2003 (revised)
 DT 25-AUG-1999 (first entry)
 XX DE Porphyromonas gingivalis protein PG67.
 XX KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.
 XX OS Porphyromonas gingivalis.
 XX PN WO9929870-A1.
 XX PD 17-JUN-1999.
 XX PF 10-DEC-1998; 98WO-AU001023.
 XX PR 10-DEC-1997; 97AU-00000839.
 PR 31-DEC-1997; 97AU-00001182.
 PR 30-JAN-1998; 98AU-00001546.
 PR 10-MAR-1998; 98AU-00002264.
 PR 09-APR-1998; 98AU-00002911.
 PR 23-APR-1998; 98AU-00003128.
 PR 05-MAY-1998; 98AU-00003338.
 PR 22-MAY-1998; 98AU-00003654.
 PR 29-JUL-1998; 98AU-00004917.
 PR 30-JUL-1998; 98AU-00004963.
 PR 04-AUG-1998; 98AU-00005028.
 XX PA (CSLC-) CSL LTD.
 XX PI Ross BC, Barr IG, Patterson MA, Agius CT, Rothel LJ, Margetts MB;
 PI Hocking DM, Webb EA;
 XX WPI; 1999-385613/32.
 DR N-PSDB; AAX91621.
 XX PT Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
 XX Claim 1; Page 373-375; 588pp; English.
 XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation


```
SQ Sequence 111 AA;
Query Match 47.1%; Score 41; DB 4; Length 111;
Best Local Similarity 38.5%; Pred. No. 60;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 13
   |::|||::|
Db 51 RGRHDIRGFAQW 63

RESULT 37
ADF05142
ID ADF05142 standard; protein; 144 AA.
XX AC ADF05142;
XX DT 12-FEB-2004 (first entry)
XX DE Bacterial polypeptide #1255.
XX KW Proteus mirabilis infection; bacterial infection; antibacterial;
XX KW immunostimulant.
XX OS Proteus mirabilis.
XX PN US6605709-B1.
XX PD 12-AUG-2003.
XX PF 05-APR-2000; 2000US-00543681.
XX PR 09-APR-1999; 99US-0128706P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL;
XX DR WPI; 2003-895291/82.
XX DR N-PSDB; ADF00970.
XX DT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX PS Disclosure; SEQ ID NO 5427; 870pp; English.
XX CC The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX SQ Sequence 144 AA;
Query Match 47.1%; Score 41; DB 7; Length 144;
Best Local Similarity 87.5%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YQHDRLAY 10
   |||||
Db 56 YQHDRLAY 63

RESULT 38
SQ Sequence 227 AA;
Query Match 47.1%; Score 41; DB 7; Length 227;
Best Local Similarity 63.6%; Pred. No. 13e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 HDLRAYGFWRL 15
   |||||
Db 14 HDAPAYGLWLL 24

RESULT 39
ADI27244
ID ADI27244 standard; protein; 497 AA.
XX AC ADI27244;
XX DT 22-APR-2004 (first entry)
XX DE B. napus BnAOX-1 amino acid sequence, SEQ ID 5.
XX
```

```
ABO76461
ID ABO76461 standard; protein; 227 AA.
XX AC ABO76461;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #8636.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI; 2003-615309/58.
XX DR N-PSDB; ABD10032.
XX DT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 25207; 455pp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 227 AA;
Query Match 47.1%; Score 41; DB 7; Length 227;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 HDLRAYGFWRL 15
   |||||
Db 14 HDAPAYGLWLL 24

RESULT 39
ADI27244
ID ADI27244 standard; protein; 497 AA.
XX AC ADI27244;
XX DT 22-APR-2004 (first entry)
XX DE B. napus BnAOX-1 amino acid sequence, SEQ ID 5.
XX
```

KW Plant; protectant; amine oxidase; stress response;
KW stress-related polypeptide; AOXSRP; environmental stress; high salinity;
KW drought; low temperature; lodging.
XX Brassica napus.
XX WO2003039243-A2.
XX 15-MAY-2003.
XX 12-NOV-2002; 2002WO-US036507.
XX 09-NOV-2001; 2001US-0345969P.
XX (BADI) BASP PLANT SCI GMBH.
XX Henkes S, Chen R, Da Costa E SilvaO, Van Thielen N;
XX WPI; 2003-430616/40.
XX N-PSDB; ADI27243.
XX New amine oxidase stress-related polypeptides and nucleic acids encoding
XX them, useful for increasing a plant's tolerance to environmental stress,
XX including high salinity, drought, low temperature and lodging.
XX Claim 1; Page 71; 77pp; English.
XX The invention relates to an isolated nucleic acid comprising a
XX polynucleotide encoding a polypeptide that functions as a modulator of a
XX plant stress response. Also disclosed as new are amino oxidase stress-
XX related polypeptide (AOXSRP) chimeric or fusion polypeptides comprising
XX an AOXSRP fused to a non-AOXSRP polypeptide, and antibodies that
XX specifically bind to an AOXSRP. The amine oxidase stress-related
XX polypeptides are useful for increasing a plant's tolerance to
XX environmental stress, including high salinity, drought, low temperature
XX and lodging, or for transforming plants tolerant to these environmental
XX stresses. The nucleic acid molecules, polypeptides, peptide homologs,
XX fusion polypeptides, primers, vectors and host cells are useful in the
XX identification of Physcomitrella patens, Brassica napus and other related
XX organisms. They are useful for the mapping of genomes of organisms
XX related to P. patens or B. napus, the identification and localisation of
XX P. patens or B. napus sequences of interest, the determination of AOXSRP
XX regions required for function, the modulation of AOXSRP activity, the
XX modulation of expression of an AOXSRP coding nucleic acid, the modulation
XX of metabolism or one or more cell functions, the modulation of
XX transmembrane transport of one or more compounds, or the modulation of
XX stress resistance. The current sequence represents the B. napus BnAOX-1
XX amino acid sequence.
XX SQ Sequence 497 AA;
Query Match 47.1%; Score 41; DB 7; Length 497;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 YQHDLRAYGFW 13
| | | | : | | :
Db 115 YQHDLESYGLY 125
RESULT 40
ADG98169
ID ADG98169 standard; protein; 544 AA.
XX AC ADG98169;
XX 11-MAR-2004 (first entry)
XX BsaI restriction endonuclease (bsaI) polypeptide.
XX BsaI restriction endonuclease; endonuclease; methylase;
KW restriction enzyme; enzyme purification; bsaI; enzyme.
XX

OS Geobacillus stearothermophilus.
XX US2003186363-A1.
XX 02-OCT-2003.
XX 26-MAR-2002; 2002US-00106275.
XX 26-MAR-2002; 2002US-00106275.
XX (NEWE) NEW ENGLAND BIOLABS INC.
XX Zhu Z, Xu S;
XX WPI; 2004-032006/03.
XX N-PSDB; ADG98168.
XX New isolated DNA obtainable from Bacillus stearothermophilus used in
XX vectors for transforming a host cell and producing recombinant BsaI
XX restriction endonuclease.
XX Disclosure; SEQ ID NO 6; 29pp; English.
XX The invention relates to DNA encoding the BsaI restriction endonuclease,
XX where the DNA is obtainable from Bacillus stearothermophilus 6-55. The
XX invention also relates to a recombinant DNA vector into which a DNA
XX segment encoding the BsaI restriction endonuclease gene has been
XX inserted, and a method for producing recombinant BsaI restriction
XX endonuclease comprising culturing a host cell transformed with the vector
XX under conditions suitable for expression of endonuclease and methylase.
XX The invention is used to produce large quantities of restriction enzymes,
XX thus simplifying enzyme purification. This sequence represents the BsaI
XX restriction endonuclease polypeptide of the invention.
XX SQ Sequence 544 AA;
Query Match 47.1%; Score 41; DB 8; Length 544;
Best Local Similarity 53.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 RSYQHDLRAYGFW 13
| | | | : | | :
Db 229 RSYTTDRRAPEYV 241
Search completed: September 16, 2005, 12:41:00
Job time : 171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 12:35:12 ; Search time 43 Seconds

(without alignments)
26.040 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSQVHDLRAYGFWRLL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	51.7	903	4	US-09-543-681A-4186
2	43	49.4	465	4	US-09-957-005-11
3	43	49.4	484	4	US-09-489-039A-9895
4	43	49.4	530	4	US-09-966-997-13
5	41	47.1	92	4	US-09-270-767-38503
6	41	47.1	92	4	US-09-270-767-53720
7	41	47.1	144	4	US-09-543-681A-5427
8	41	47.1	227	4	US-09-252-991A-25207
9	41	47.1	544	4	US-10-106-275-6
10	40	46.0	151	4	US-09-765-815-13
11	40	46.0	334	4	US-09-287-599A-2
12	40	46.0	920	4	US-09-934-868-52
13	39	44.8	239	4	US-09-248-796A-22464
14	39	44.8	265	5	PCR-US95-03776-29
15	39	44.8	265	5	US-08-413-803-27
16	39	44.8	286	1	US-08-321-488A-27
17	39	44.8	433	4	US-09-711-164-412
18	39	44.8	504	4	US-09-248-796A-17765
19	39	44.8	2104	2	US-08-808-793-4
20	39	44.8	2104	3	US-08-772-512A-4
21	39	44.8	2104	4	US-09-428-371-4
22	39	44.8	2105	2	US-08-808-793-3
23	39	44.8	2105	3	US-08-772-512A-3
24	39	44.8	2105	4	US-09-428-371-3
25	38	43.7	66	3	US-08-936-165A-376
26	38	43.7	109	4	US-09-328-352-7204
27	38	43.7	148	3	US-08-946-329A-65
28	38	43.7	181	4	US-09-107-532A-4183
29	38	43.7	320	1	US-08-362-739-2
30	38	43.7	320	4	US-08-914-350A-2
31	38	43.7	557	4	US-09-902-540-16548
32	38	43.7	681	4	US-09-252-991A-28831
33	38	43.7	827	4	US-09-902-540-13652
34	38	43.7	1111	1	US-08-317-450B-15
35	38	43.7	1111	3	US-08-800-593-15
36	38	43.7	1171	1	US-08-445-135-1
37	38	43.7	1171	4	US-09-560-385A-36
38	38	43.7	1172	4	US-09-560-385A-28
39	38	43.7	1172	4	US-09-560-385A-32
40	38	43.7	1192	4	US-09-560-385A-34
41	38	43.7	1193	1	US-08-317-450B-13
42	38	43.7	1193	3	US-08-800-593-13
43	38	43.7	1193	4	US-09-560-385A-26
44	38	43.7	1193	4	US-09-560-385A-30
45	37	42.5	16	4	US-09-344-783C-2
46	37	42.5	16	4	US-09-069-827A-115
47	37	42.5	76	3	US-09-227-357-456
48	37	42.5	187	3	US-09-247-155-102
49	37	42.5	190	4	US-09-270-767-33036
50	37	42.5	190	4	US-09-270-767-48253
51	37	42.5	290	4	US-09-252-991A-17890
52	37	42.5	489	4	US-09-807-258-29
53	37	42.5	567	2	US-08-504-459-2
54	37	42.5	567	2	US-08-504-459-6
55	37	42.5	616	1	US-08-149-100-2
56	37	42.5	629	4	US-09-902-540-12912
57	37	42.5	661	4	US-09-328-352-6748
58	37	42.5	695	6	5460961-5
59	37	42.5	695	6	5460961-5
60	37	42.5	761	4	US-09-949-016-9802
61	37	42.5	772	4	US-09-252-991A-30121
62	37	42.5	877	1	US-08-430-925A-4
63	37	42.5	940	2	US-08-938-365-4
64	37	42.5	941	1	US-08-343-760A-2
65	37	42.5	1770	4	US-09-902-540-10224
66	36.5	42.0	314	4	US-09-489-039A-9630
67	36.5	42.0	326	4	US-09-252-991A-30438
68	36.5	42.0	352	4	US-09-328-352-4999
69	36	41.4	74	2	US-08-743-130A-41
70	36	41.4	189	4	US-09-134-000C-5612
71	36	41.4	209	4	US-09-328-352-7988
72	36	41.4	261	4	US-09-270-767-47059
73	36	41.4	291	4	US-09-328-352-4557
74	36	41.4	310	4	US-09-248-796A-17322
75	36	41.4	319	4	US-09-902-540-14497
76	36	41.4	332	4	US-09-198-452A-375
77	36	41.4	374	4	US-09-721-870-24
78	36	41.4	592	4	US-09-813-453B-22
79	36	41.4	592	4	US-09-813-453B-43
80	36	41.4	682	4	US-09-107-532A-6403
81	36	41.4	749	4	US-09-562-737-97
82	36	41.4	896	4	US-09-328-352-6776
83	36	41.4	1665	4	US-09-543-681A-4476
84	35.5	40.8	67	2	US-08-781-891-75
85	35.5	40.8	211	4	US-09-618-166-75
86	35.5	40.8	249	4	US-09-710-279-420
87	35.5	40.8	252	4	US-09-328-352-6990
88	35.5	40.8	282	4	US-08-311-731A-271
89	35.5	40.8	369	4	US-09-107-532A-6403
90	35.5	40.8	607	3	US-09-543-681A-17173
91	35.5	40.8	607	4	US-09-618-166-75
92	35.5	40.8	766	4	US-09-710-279-420
93	35.5	40.8	2323	4	US-09-710-279-760
94	35.5	40.8	2404	3	US-09-134-001C-3464
95	35	40.2	53	3	US-09-187-789-65
96	35	40.2	53	4	US-09-139-600-60
97	35	40.2	53	4	US-09-989-903-65
98	35	40.2	114	4	US-09-087-031E-15
99	35	40.2	115	4	US-09-087-031E-10
100	35	40.2	119	4	US-09-107-532A-4937

Sequence 4183, Ap
Sequence 2, Appli
Sequence 2, Appli
Sequence 16548, A
Sequence 28831, A
Sequence 13652, A
Sequence 15, Appl
Sequence 15, Appl
Sequence 1, Appli
Sequence 36, Appl
Sequence 28, Appl
Sequence 32, Appl
Sequence 34, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 26, Appl
Sequence 30, Appl
Sequence 2, Appli
Sequence 115, App
Sequence 456, App
Sequence 102, App
Sequence 33036, A
Sequence 48253, A
Sequence 17890, A
Sequence 29, Appl
Sequence 2, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 12912, A
Sequence 6748, Ap
Patent No. 5460961
Sequence 9802, Ap
Sequence 30121, A
Sequence 4, Appli
Sequence 2, Appli
Sequence 10224, A
Sequence 9630, Ap
Sequence 30438, A
Sequence 4959, Ap
Sequence 41, Appl
Sequence 5612, Ap
Sequence 7988, Ap
Sequence 47059, A
Sequence 4557, Ap
Sequence 17322, A
Sequence 14497, A
Sequence 375, App
Sequence 24, Appl
Sequence 22, Appl
Sequence 43, Appl
Sequence 6403, Ap
Sequence 97, Appl
Sequence 6776, Ap
Sequence 4776, Ap
Sequence 100, App
Sequence 6950, Ap
Sequence 1032, Ap
Sequence 5944, Ap
Sequence 7173, Ap
Sequence 75, Appl
Sequence 75, Appl
Sequence 420, App
Sequence 760, App
Sequence 65, Appl
Sequence 60, Appl
Sequence 65, Appl
Sequence 15, Appl
Sequence 1037, Ap
Sequence 4937, Ap

ALIGNMENTS

```
RESULT 1
US-09-543-681A-4186
; Sequence 4186, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4186
; LENGTH: 903
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4186

Query Match      51.7%; Score 45; DB 4; Length 903;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 SYQHDLRAYGFW 14
Db      110 NYRHRDLRGYFW 122
      :|||:|||||
      :|||:|||||

RESULT 2
US-09-957-005-11
; Sequence 11, Application US/09957005
; Patent No. 6596524
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenyu
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsmAI Restriction
; FILE REFERENCE: NEB-192
; CURRENT APPLICATION NUMBER: US/09/957,005
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus A664
US-09-957-005-11

Query Match      49.4%; Score 43; DB 4; Length 465;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDLRAYGFW 13
Db      214 KSYTKDRRAYEYW 226
      :|||:|||||:|
      :|||:|||||:|

RESULT 3
US-09-489-039A-9895
; Sequence 9895, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
```

```
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9895
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9895

Query Match      49.4%; Score 43; DB 4; Length 484;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      8 RAYGFWRL 15
Db      448 RRYGFWRL 455
      :|||||
      :|||||

RESULT 4
US-09-966-997-13
; Sequence 13, Application US/09966997
; Patent No. 6764843
; GENERAL INFORMATION:
; APPLICANT: Xu, Shuang-yong
; APPLICANT: Dore, Andrew
; APPLICANT: Hume, Adam
; APPLICANT: Pelletier, John
; APPLICANT: Zhou, Jing
; TITLE OF INVENTION: Method For Cloning And Expression of BsmBI Restriction
; TITLE OF INVENTION: Endonuclease And BsmBI Methylase In E. coli And
; FILE REFERENCE: NEB-191
; CURRENT APPLICATION NUMBER: US/09/966,997
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus B61
US-09-966-997-13

Query Match      49.4%; Score 43; DB 4; Length 530;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDLRAYGFW 13
Db      226 RLYNHDRRAFWMW 238
      :|||:|||||:|
      :|||:|||||:|

RESULT 5
US-09-270-767-38503
; Sequence 38503, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38503
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-38503

Query Match      47.1%; Score 41; DB 4; Length 92;
Best Local Similarity 87.5%; Pred. No. 8.1;
```

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 DLRAYGFW 13
|||||
Db 45 DLRAYPFW 52

RESULT 6

US-09-270-767-53720
; Sequence 53720, Application US/09270767
; Patent No. 6703491

GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7325-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53720
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53720

Query Match 47.1%; Score 41; DB 4; Length 92;

Best Local Similarity 87.5%; Pred. No. 8.1; Mismatches 0; Indels 1; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 DLRAYGFW 13
|||||
Db 45 DLRAYPFW 52

RESULT 7

US-09-543-681A-5427
; Sequence 5427, Application US/09543681A
; Patent No. 6605709

GENERAL INFORMATION:

; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5427
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5427

Query Match 47.1%; Score 41; DB 4; Length 144;

Best Local Similarity 87.5%; Pred. No. 13; Mismatches 0; Indels 1; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YQHDRLAY 10
|||||
Db 56 YQHDRLAY 63

RESULT 8

US-09-252-991A-25207
; Sequence 25207, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25207
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25207

Query Match 47.1%; Score 41; DB 4; Length 227;

Best Local Similarity 63.6%; Pred. No. 21; Mismatches 0; Indels 4; Gaps 0;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 HDLRAYGFWL 15
|||||
Db 14 HDAPAYGLWLL 24

RESULT 9

US-10-106-275-6
; Sequence 6, Application US/10106275
; Patent No. 6723546

GENERAL INFORMATION:

; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of Bsal Restriction Endonuclease
; FILE REFERENCE: NEB-202
; CURRENT APPLICATION NUMBER: US/10/106,275
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus 6-55
US-10-106-275-6

Query Match 47.1%; Score 41; DB 4; Length 544;

Best Local Similarity 53.8%; Pred. No. 55; Mismatches 2; Indels 4; Gaps 0;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLAYGFW 13
|||||
Db 229 RSYTDRRAPEYW 241

RESULT 10

US-09-765-815-13
; Sequence 13, Application US/09765815
; Patent No. 6673586

GENERAL INFORMATION:

; APPLICANT: Balk, Steven
; TITLE OF INVENTION: No. 6673586el Steroid Hormone Receptor
; FILE REFERENCE: 01948/068002
; CURRENT APPLICATION NUMBER: US/09/765,815
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/176,859
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-765-815-13

Query Match 46.0%; Score 40; DB 4; Length 151;

Best Local Similarity 50.0%; Pred. No. 21;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGCF 12
| | | | | : | |
Db 52 RYQHEVEMYNF 63

RESULT 11
US-09-287-599A-2
; Sequence 2, Application US/09287599A
; Patent No. 6602712
; GENERAL INFORMATION:
; APPLICANT: Handelsman, Jo
; APPLICANT: Klimowicz, Amy K
; TITLE OF INVENTION: Enterotoxin-Deficient Bacillus
; FILE REFERENCE: 960296.95327
; CURRENT APPLICATION NUMBER: US/09/287,599A
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/080943
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Bacillus
; OTHER INFORMATION: cereus hbla, having 123 bp in-frame deletion
; OTHER INFORMATION: spanning nucleotides 568-690
US-09-287-599A-2

Query Match 46.0%; Score 40; DB 4; Length 334;
Best Local Similarity 63.6%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGCF 12
| | | | | : | |
Db 181 SIGHDVRAFGF 191

RESULT 12
US-09-934-868-52
; Sequence 52, Application US/09934868
; Patent No. 6689601
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odcom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C11596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 52
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Methylomonas 16a
; FEATURE:
; OTHER INFORMATION: Nasa
US-09-934-868-52

Query Match 46.0%; Score 40; DB 4; Length 920;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWR 14
| | | | | : | |
Db 727 YHVKGHGFWR 738

RESULT 13
US-09-248-796A-22464
; Sequence 22464, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22464
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22464

Query Match 44.8%; Score 39; DB 4; Length 239;
Best Local Similarity 58.3%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWR 14
| | | | | : | |
Db 2 YVHDLRKPFPR 13

RESULT 14
US-08-413-803-27
; Sequence 27, Application US/08413803
; Patent No. 5766581
; GENERAL INFORMATION:
; APPLICANT: Bartley, Timothy D.
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Hunt, Pamela
; APPLICANT: Kinstler, Olaf B.
; APPLICANT: Samal, Babru B.
; TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH
; TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE MEGAKARYOCYTE
; TITLE OF INVENTION: GROWTH AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: US
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,803
; FILING DATE: 30-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,768
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/252,628
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/321,488
; FILING DATE: 12-OCT-1994
; PRIOR APPLICATION DATA:


```

/ APPLICATION NUMBER: US 08/347,780
/ FILING DATE: 30-NOV-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Cook Ph.D., Robert R.
/ REGISTRATION NUMBER: 31,602
/ REFERENCE/DOCKET NUMBER: A-290D
/ INFORMATION FOR SEQ ID NO: 27:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 265 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-413-803-27

```

```
Query Match      44.8%; Score 39; DB 1; Length 265;
Best Local Similarity 58.3%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

Qy 2 SYQHDLRAYGFW 13
|:|:|:|:|
Db 130 SFOHLLRGKDFW 141

RESULT 15

```

PCT-US95-03776-29
; Sequence 29, Application PC/TUS9503776
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Compositions and Methods for Stimulating
; TITLE OF INVENTION: Megakaryocyte Growth and Differentiation
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03776
;

```

ATTORNEY/AGENT INFORMATION:
NAME: Cook, Robert R.
REFERENCES/DOCKET NUMBER: A-290-C
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 265 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PRT-UG95-03776-29

Query Match 44.8%; Score 39; DB 5; Length 265;
Best Local Similarity 58.3%; Pred. No. 56;
Matches 7: Conservative 1; Mismatches 4; Indels

Qy 2 SYQHDLRAYGFW 13
|:|:|:|:|
Db 130 SFOHLLRGKDFW 141

RESULT 16
US-08-321-488A-27
; Sequence 27, Application US/08321488A
; Patent No. 5795569
; GENERAL INFORMATION:
; APPLICANT: Bartley, Timothy D.

```

1 / APPLICANT: Bogenberger, Jakob M.
2 / APPLICANT: Bosseelman, Robert A.
3 / APPLICANT: Hunt, Pamela
4 / APPLICANT: Kinstler, Olaf B.
5 / APPLICANT: Samal, Babru B.
6 / TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE
7 / TITLE OF INVENTION: MEGAKARYOCYTE GROWTH AND DIFFERENTIATION
8 / NUMBER OF SEQUENCES: 27
9 / CORRESPONDENCE ADDRESS:
10 / ADDRESSEE: AMGEN INC.
11 / STREET: 1840 DeHavilland Drive
12 / CITY: Thousand Oaks
13 / STATE: California
14 / COUNTRY: US
15 / ZIP: 91320-1789
16 / COMPUTER READABLE FORM:
17 / MEDIUM TYPE: Floppy disk
18 / COMPUTER: IBM PC compatible
19 / OPERATING SYSTEM: PC-DOS/MS-DOS
20 / SOFTWARE: PatentIn Release #1.0, Version #1.30
21 / CURRENT APPLICATION DATA:
22 / APPLICATION NUMBER: US/08/321.488A
23 / FILING DATE: 12-OCT-1994
24 / CLASSIFICATION: 435
25 / PRIOR APPLICATION DATA:
26 / APPLICATION NUMBER: US 08/252,628
27 / FILING DATE: 31-MAY-1994
28 / PRIOR APPLICATION DATA:
29 / APPLICATION NUMBER: US 08/221,768
30 / FILING DATE: 31-MAR-1994
31 / ATTORNEY/AGENT INFORMATION:
32 / NAME: Cook, Robert R.
33 / REGISTRATION NUMBER: 31,602
34 / REFERENCE/DOCKET NUMBER: A-2908
35 / INFORMATION FOR SEQ ID NO: 27:
36 / SEQUENCE CHARACTERISTICS:
37 / LENGTH: 286 amino acids
38 / TYPE: amino acid
39 / TOPOLOGY: linear
40 / MOLECULE TYPE: protein
41 / PS-08-321-488A-27

```

Query Match 44.8%; Score 39; DB 1; Length 286;
Best Local Similarity 58.3%; Pred. No. 61;
Matches 7: Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy      2 SYQHDLRAYGEW 13
        | : | | | |
Db     151 SFOHLLRGKDEW 162
```

```

RESULT 17
US-09-711-164-412
; Sequence 412, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Orlsyth, R. Aliyn
; APPLICANT: Orlsenth, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR M
; FILE REFERENCE: ELTRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-412

```

Query Match 44.8%; Score 39; DB 4; Length 433;
Best Local Similarity 46.2%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWR 15
DB 410 FWHHLRRHGLWHL 422

RESULT 18
US-09-248-796A-17765
; Sequence 17765, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17765
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (481)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-17765

Query Match 44.8%; Score 39; DB 4; Length 504;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 HDLRAYGFWR 15
DB 166 HDERDFDWKL 176

RESULT 19
US-08-808-793-4
; Sequence 4, Application US/08808793
; Patent No. 5858713
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,793
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/034,361
; FILING DATE: 24-DEC-1996

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,649
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brame, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2104 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-808-793-4

Query Match 44.8%; Score 39; DB 2; Length 2104;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWR 15
DB 383 YQHVLAAGPWHM 395

RESULT 20
US-08-772-512A-4
; Sequence 4, Application US/08772512A
; Patent No. 6022705
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,512A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,618
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brame, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/601 (CRFD-1657)
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2104 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-772-512A-4

Query Match 44.8%; Score 39; DB 3; Length 2104;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHLDRAYGFWRLL 15
Db 383 YQHVLAAGPWHM 395

RESULT 21
US-09-428-371-4
; Sequence 4, Application US/09428371
; Patent No. 6800435
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM INSECTICIDE-RESISTANT HOUSE FLIES
; FILE REFERENCE: 19603/606
; CURRENT APPLICATION NUMBER: US/09/428,371
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 08/608,618
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: 08/772,512
; EARLIER FILING DATE: 1996-12-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2104
; TYPE: PRT
; ORGANISM: Musca domestica
US-09-428-371-4

Query Match 44.8%; Score 39; DB 4; Length 2104;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHLDRAYGFWRLL 15
Db 383 YQHVLAAGPWHM 395

RESULT 22
US-08-808-793-3
; Sequence 3, Application US/08808793
; Patent No. 5858713
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
; NUMBER OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/808,793
; CLASSIFICATION: 435
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/034,361
; FILING DATE: 24-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US 60/012,649
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Braman, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 2105 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-808-793-3

Query Match 44.8%; Score 39; DB 2; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHLDRAYGFWRLL 15
Db 383 YQHVLAAGPWHM 395

RESULT 23
US-08-772-512A-3
; Sequence 3, Application US/08772512A
; Patent No. 6022705
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
; TITLE OF INVENTION: FLIES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/772,512A
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,618
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Braman, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/601 (CRFD-1657)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 2105 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-772-512A-3

Query Match          44.8%; Score 39; DB 3; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 YQHDRLRAYGFWR 15
DB      383 YQHVLAAGPWHM 395

RESULT 24
US-09-428-371-3
; Sequence 3, Application US/09428371
; Patent No. 6800435
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM INSECTICIDE-SUSCEPTIBLE AND
; TITLE OF INVENTION: INSECTICIDE-RESISTANT HOUSE FLIES
; FILE REFERENCE: 19603/606
; CURRENT APPLICATION NUMBER: US/09/428,371
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 08/608,618
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: 08/772,512
; EARLIER FILING DATE: 1996-12-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2105
; TYPE: PRT
; ORGANISM: Musca domestica
US-09-428-371-3

Query Match          44.8%; Score 39; DB 4; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 YQHDRLRAYGFWR 15
DB      383 YQHVLAAGPWHM 395

RESULT 25
US-08-936-165A-376
; Sequence 376, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

; LENGTH: 2105 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-772-512A-3

Query Match          44.8%; Score 39; DB 3; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 YQHDRLRAYGFWR 15
DB      383 YQHVLAAGPWHM 395

RESULT 24
US-09-428-371-3
; Sequence 3, Application US/09428371
; Patent No. 6800435
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM INSECTICIDE-SUSCEPTIBLE AND
; TITLE OF INVENTION: INSECTICIDE-RESISTANT HOUSE FLIES
; FILE REFERENCE: 19603/606
; CURRENT APPLICATION NUMBER: US/09/428,371
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 08/608,618
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: 08/772,512
; EARLIER FILING DATE: 1996-12-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2105
; TYPE: PRT
; ORGANISM: Musca domestica
US-09-428-371-3

Query Match          44.8%; Score 39; DB 4; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 YQHDRLRAYGFWR 15
DB      383 YQHVLAAGPWHM 395

RESULT 25
US-08-936-165A-376
; Sequence 376, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimml, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 376:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-936-165A-376

Query Match          43.7%; Score 38; DB 3; Length 66;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 SYQHDRLRAYGFWR 15
DB      25 SYEPPIEAGGTWRL 38

RESULT 26
US-09-328-352-7204
; Sequence 7204, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7204
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7204

Query Match          43.7%; Score 38; DB 4; Length 109;
Best Local Similarity 42.9%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 RSYQHDRLRAYGFWR 14
DB      41 KAYSQDLORQGWKR 54

RESULT 27
US-08-946-329A-65
; Sequence 65, Application US/08946329A
; Patent No. 6057091
; GENERAL INFORMATION:
; APPLICANT: Beachy, Philip A.
```

APPLICANT: Porter, Jeffrey A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946.329A
FILING DATE: 07-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 148 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-946-329A-65

Query Match 43.7%; Score 38; DB 3; Length 148;
Best Local Similarity 53.8%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGFWR 14
Db 105 SQDHDLEGRGKWR 117

RESULT 28
US-09-107-532A-4183
Sequence 4183, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4183:
SEQUENCE CHARACTERISTICS:
LENGTH: 181 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...181
SEQUENCE DESCRIPTION: SEQ ID NO: 4183:
US-09-107-532A-4183

Query Match 43.7%; Score 38; DB 4; Length 181;
Best Local Similarity 42.9%; Pred. No. 55;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSQHDRLRAYGFWR 14
Db 17 QSYKHDKGLHRTWR 30

RESULT 29
US-08-362-739-2
Sequence 2, Application US/08362739
Patent No. 5658757
GENERAL INFORMATION:
APPLICANT: Haake, David A.
APPLICANT: Blanco, David R.
APPLICANT: Champion, Cheryl I.
APPLICANT: Lovett, Michael A.
APPLICANT: Miller, James N.
TITLE OF INVENTION: CLONED Leptospira OUTER MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,739
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/040,747
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John R.,
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-2097

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-739-2

Query Match 43.7%; Score 38; DB 1; Length 320;
Best Local Similarity 46.2%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWR 15
| : | | | | |
Db 131 YESDFGKYFWRV 143

RESULT 30

US-08-914-350A-2
; Sequence 2, Application US/08914350A
; Patent No. 6685945
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

; APPLICANT: HAAKE, David A.
; APPLICANT: BLANCO, David R.
; APPLICANT: CHAMPION, Cheryl I.
; APPLICANT: LOVETT, Michael A.
; APPLICANT: MILLER, James N.
; TITLE OF INVENTION: CLONED LEPTOSPIRA OUTER MEMBRANE PROTEIN
; FILE REFERENCE: UCL1420-2
; CURRENT APPLICATION NUMBER: US/08/914,350A
; CURRENT FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 08/362,739
; PRIOR FILING DATE: 1994-12-20
; PRIOR APPLICATION NUMBER: US 08/040,747
; PRIOR FILING DATE: 1993-03-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2

; LENGTH: 320

; TYPE: PRT

; ORGANISM: Leptospira kirschneri
; US-08-914-350A-2

Query Match 43.7%; Score 38; DB 4; Length 320;
Best Local Similarity 46.2%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWR 15
| : | | | | |
Db 131 YESDFGKYFWRV 143

RESULT 31

US-09-902-540-16548
; Sequence 16548, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16548

; LENGTH: 557
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-16548

Query Match 43.7%; Score 38; DB 4; Length 557;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 DLRAYGFWR 14
| | | | |
Db 334 DLLARGFWR 342

RESULT 32

US-09-252-991A-28831
; Sequence 28831, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28831
; LENGTH: 681

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-28831

Query Match 43.7%; Score 38; DB 4; Length 681;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGF 12
| : | | | | |
Db 440 REFAEDIRMYGF 451

RESULT 33

US-09-902-540-13652
; Sequence 13652, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13652
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-13652

Query Match 43.7%; Score 38; DB 4; Length 827;
Best Local Similarity 60.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 QHDLRAYGFWR 13
| | | | |

Db 644 ERKLAYGLW 653

RESULT 34

US-08-317-450B-15

Sequence 15, Application US/08317450B

Patent No. 5660982

GENERAL INFORMATION:

APPLICANT: Tryggvason, Karl

APPLICANT: Kallunki, Pekka

APPLICANT: Pyke, Charles

TITLE OF INVENTION: Laminin Chains: Diagnostic and

TITLE OF INVENTION: Therapeutic Use

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: BANNER & ALLEGRETTI, LTD.

STREET: Ten South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/317.450B

FILING DATE: 04-OCT-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Chao, Mark

REGISTRATION NUMBER: 37,293

REFERENCE/DOCKET NUMBER: 94,778

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1234

TELEFAX: 312-715-1234

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1111 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-317-450B-15

Query Match 43.7%; Score 38; DB 1; Length 1111;

Best Local Similarity 87.5%; Pred. No. 3.9e+02;

Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSYQHDLR 8

Db 843 RSYQHSLR 850

RESULT 35

US-08-800-593-15

Sequence 15, Application US/08800593

Patent No. 6143505

GENERAL INFORMATION:

APPLICANT: Tryggvason, Karl

APPLICANT: Kallunki, Pekka

APPLICANT: Pyke, Charles

TITLE OF INVENTION: Laminin Chains: Diagnostic and

TITLE OF INVENTION: Therapeutic Use

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/445.135

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 08/151,134

FILING DATE: 12-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: DESMOS.002CP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 1:

Query Match 43.7%; Score 38; DB 3; Length 1111;

Best Local Similarity 87.5%; Pred. No. 3.9e+02;

Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSYQHDLR 8

Db 843 RSYQHSLR 850

RESULT 36

US-08-445-135-1

Sequence 1, Application US/08445135

Patent No. 5658789

GENERAL INFORMATION:

APPLICANT: Quaranta, Vito

APPLICANT: Hormia, Marketta

TITLE OF INVENTION: Promotion of Epithelial Cell Adhesion

TITLE OF INVENTION: and Hemidesmosome Assembly by a Laminin-Like Molecule

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson and Bear

STREET: 620 Newport Center Drive, Sixteenth Floor

CITY: Newport Beach

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/445.135

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 08/151,134

FILING DATE: 12-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: DESMOS.002CP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 1:

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1171 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FRAGMENT TYPE: internal
; US-08-445-135-1
;
Query Match      43.7%; Score 38; DB 1; Length 1171;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RSYQHDLR 8
Db      822 RSYQHSRLR 829

RESULT 37
US-09-560-385A-36
; Sequence 36, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 1171
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-560-385A-36

Query Match      43.7%; Score 38; DB 4; Length 1171;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RSYQHDLR 8
Db      823 RSYQHSRLR 830

RESULT 38
US-09-560-385A-28
; Sequence 28, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-560-385A-28

Query Match      43.7%; Score 38; DB 4; Length 1172;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RSYQHDLR 8
Db      822 RSYQHSRLR 829
```

```
RESULT 39
US-09-560-385A-32
; Sequence 32, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-560-385A-32

Query Match      43.7%; Score 38; DB 4; Length 1172;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RSYQHDLR 8
Db      822 RSYQHSRLR 829

RESULT 40
US-09-560-385A-34
; Sequence 34, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-560-385A-34

Query Match      43.7%; Score 38; DB 4; Length 1192;
Best Local Similarity 87.5%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RSYQHDLR 8
Db      844 RSYQHSRLR 851

Search completed: September 16, 2005, 12:45:25
Job time : 45 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 12:44:03 ; Search time 166 Seconds
(without alignment)
36.590 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHLRLAYCFWRL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	56.3	554	16	US-10-437-963-169736
2	47	54.0	554	18	US-10-450-763-50090
3	47	54.0	554	18	US-10-450-763-50144
4	47	54.0	1007	18	US-10-450-763-33442
5	44	50.6	267	18	US-10-450-763-56548
6	44	50.6	1356	15	US-10-137-824-35
7	44	50.6	1593	15	US-10-197-824-19
8	44	50.6	1891	18	US-10-450-763-34690
9	43	49.4	29	19	US-11-028-539-34
10	43	49.4	256	14	US-10-156-761-9242
11	43	49.4	360	19	US-11-028-539-36

43	49.4	420	19	US-11-028-539-58	Sequence 58, Appl
43	49.4	449	20	US-11-013-235-38	Sequence 18, Appl
43	49.4	465	15	US-09-957-005-11	Sequence 11, Appl
43	49.4	466	15	US-10-282-122A-59950	Sequence 59950, A
43	49.4	525	20	US-11-013-235-32	Sequence 32, Appl
43	49.4	530	20	US-09-966-997-13	Sequence 13, Appl
43	49.4	530	20	US-10-425-114-59634	Sequence 59634, A
42.5	48.9	413	15	US-10-425-115-194035	Sequence 194035, A
42.5	48.9	461	16	US-10-425-115-233513	Sequence 233513, A
42	48.3	370	16	US-10-437-963-194505	Sequence 194505, A
42	48.3	370	16	US-10-437-963-194505	Sequence 42825, A
42	48.3	558	16	US-10-767-701-42825	Sequence 121169, A
42	48.3	891	16	US-10-437-963-121169	Sequence 136, App
42	48.3	994	18	US-10-972-963-136	Sequence 246608, A
41	47.1	63	16	US-10-425-115-246608	Sequence 348393, A
41	47.1	68	16	US-10-425-115-348393	Sequence 41087, A
41	47.1	111	18	US-10-450-763-41087	Sequence 56326, A
41	47.1	111	18	US-10-450-763-56326	Sequence 253060, A
41	47.1	127	16	US-10-425-115-253060	Sequence 211706, A
41	47.1	129	15	US-10-424-599-211706	Sequence 4075, Ap
41	47.1	301	17	US-10-732-923-4075	Sequence 37, Appl
41	47.1	470	20	US-11-013-235-37	Sequence 5, Appl
41	47.1	497	14	US-10-293-958-5	Sequence 31, Appl
41	47.1	524	20	US-11-013-235-31	Sequence 6, Appl
41	47.1	544	14	US-10-106-273-6	Sequence 129380, A
41	47.1	573	16	US-10-437-963-129380	Sequence 578, App
41	47.1	588	18	US-10-220-335-578	Sequence 38718, A
41	47.1	651	18	US-10-450-763-38718	Sequence 1468, Ap
41	47.1	692	15	US-10-389-566-1468	Sequence 1481, Ap
41	47.1	817	15	US-10-389-566-1481	Sequence 217364, A
41	47.1	1081	16	US-10-425-115-217364	Sequence 217358, A
41	47.1	1339	16	US-10-425-115-217358	Sequence 181030, A
41	47.1	1479	16	US-10-437-963-181030	Sequence 217361, A
41	47.1	1540	16	US-10-425-115-217361	Sequence 217360, A
41	47.1	1606	16	US-10-425-115-217360	Sequence 254920, A
41	47.1	1742	16	US-10-425-115-254920	Sequence 254923, A
41	47.1	1753	16	US-10-425-115-254923	Sequence 281147, A
40	46.0	76	15	US-10-424-599-281147	Sequence 38888, A
40	46.0	87	14	US-10-450-763-38888	Sequence 12936, A
40	46.0	150	14	US-10-156-761-12936	Sequence 13, Appl
40	46.0	151	10	US-09-765-815-13	Sequence 3989, Ap
40	46.0	151	10	US-10-732-923-3989	Sequence 200718, A
40	46.0	223	14	US-10-380-132-18	Sequence 3998, Ap
40	46.0	224	16	US-10-425-115-200718	Sequence 3974, Ap
40	46.0	230	17	US-10-732-923-3998	Sequence 3975, Ap
40	46.0	306	17	US-10-732-923-3974	Sequence 2, Appl
40	46.0	318	17	US-10-732-923-3975	Sequence 47984, A
40	46.0	334	9	US-09-287-599-2	Sequence 166245, A
40	46.0	478	15	US-10-425-114-47984	Sequence 166, App
40	46.0	490	15	US-10-424-599-166245	Sequence 142, App
40	46.0	618	14	US-10-295-403-166	Sequence 894, App
40	46.0	619	10	US-09-934-455-142	Sequence 164, App
40	46.0	619	15	US-10-225-066A-894	Sequence 566, App
40	46.0	619	15	US-10-374-780A-164	Sequence 894, App
40	46.0	619	15	US-10-412-699B-566	Sequence 894, App
40	46.0	619	18	US-10-225-066A-894	Sequence 40984, A
40	46.0	640	18	US-10-450-763-40984	Sequence 1551, Ap
40	46.0	655	17	US-10-732-923-1551	Sequence 183525, A
40	46.0	677	16	US-10-437-963-183525	Sequence 52, Appl
40	46.0	920	9	US-09-934-868-52	Sequence 14, Appl
40	46.0	920	11	US-09-789-161-14	Sequence 52, Appl
40	46.0	920	18	US-10-701-200-52	Sequence 274036, A
39.5	45.4	103	16	US-10-425-115-274036	Sequence 144036, A
39.5	45.4	257	16	US-10-437-963-144036	Sequence 194695, A
39	44.8	68	16	US-10-437-963-194695	Sequence 155268, A
39	44.8	72	15	US-10-424-599-165268	Sequence 175273, A
39	44.8	91	15	US-10-424-599-175273	Sequence 175274, A
39	44.8	91	16	US-10-739-930-8737	Sequence 8737, Ap
39	44.8	112	15	US-10-424-599-249688	Sequence 249688, A
39	44.8	117	15	US-10-424-599-228673	Sequence 228673, A
39	44.8	166	15	US-10-424-599-254263	Sequence 254263, A
39	44.8	198	14	US-10-204-887-96	Sequence 96, Appl
39	44.8	216	17	US-10-831-070-72	Sequence 72, Appl

Sequence 48936, A
Sequence 205777, A
Sequence 255332, A
Sequence 34867, A
Sequence 47913, A
Sequence 17199, A
Sequence 14376, A
Sequence 14569, A
Sequence 14886, A
Sequence 412, App
Sequence 42701, A
Sequence 24015, A
Sequence 194616, A
Sequence 48139, A
Sequence 15749, A
Sequence 44861, A

39	44.8	224	15	US-10-282-122A-48936
39	44.8	229	15	US-10-424-599-205777
39	44.8	241	15	US-10-424-599-255337
39	44.8	263	18	US-10-450-763-34867
39	44.8	263	18	US-10-450-763-47913
39	44.8	313	15	US-10-369-493-11799
39	44.8	313	15	US-10-369-493-14376
39	44.8	313	15	US-10-369-493-14569
39	44.8	313	15	US-10-369-493-14886
39	44.8	433	14	US-10-287-274-412
39	44.8	433	15	US-10-282-122A-42701
39	44.8	521	17	US-10-732-923-24015
39	44.8	588	16	US-10-437-963-194616
39	44.8	671	18	US-10-450-763-48139
39	44.8	699	17	US-10-732-923-15749
39	44.8	743	18	US-10-450-763-44861

ALIGNMENTS

```

RESULT 1
US-10-437-963-169736
; Sequence 169736, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 169736
; LENGTH: 554
; TYPE: PRI
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68127C.1.pgp
US-10-437-963-169736

```

Query Match	56.3%	Score 49;	DB 16;	Length 554;
Best Local Similarity	50.0%;	Pred. No. 18;		
Matches	6;	Conservative	4;	Mismatches 2;
				Indels 0;
				Gaps 0;

QY	3	YQHD	LRAY	GFWR	14
		:	:	:	:
Db	276	FRHD	IGMY	GYWR	287

```

RESULT 2
US-10-450-763-50090
; Sequence 50090, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23

```

```

; NUMBER OF SEQ ID NOS: 60736
;
; SOFTWARE: Custom
; SEQ ID NO 5090
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-450-763-50090

Query Match      54.0%   Score 47;   DB 18;   Length 554;
Best Local Similarity 53.8%   Pred. No. 38;
Matches 7;   Conservative 4;   Indels 2;   Mismatches 4;

```

Qy 3 YQHD LRAYGFWR L 15
| : | | | | :
Db 7 YKHS L HAYGAWAI 19

```

RESULT 3
US-10-450-763-50144
; Sequence 50144, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790C1P3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50144
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-50144

```

Query Match	54.0%	Score 47	DB 18	Length 554
Best Local Similarity	53.8%	Pred. No. 38		
Matches	7	Conservative	2	Mismatches
			4	Indels
			0	Gaps
			0	

Qy 3 YQHDLRAYGFWR L 15
| : | | | | :
Db 7 YKHS L HAYGAWA I 19

```

RESULT 4
US-10-450-763-33442
/ Sequence 33442, Application US/10450763
/ Publication No. US20050196754A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ PRIOR FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 33442
/ LENGTH: 1007
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:

```

; NAME/KEY: DOMAIN
; LOCATION: (939)..(950)
; OTHER INFORMATION: RNase H domain identified by eMATRIX, accession number
; OTHER INFORMATION: PF00075D, p-value=8.364e-11, raw score of 10.71
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (803)..(994)
; OTHER INFORMATION: Reverse transcriptase domain identified by PFam, accession
; OTHER INFORMATION: name rvt, E-value=1.1e-17, Pfam score of 72.2
US-10-450-763-33442

Query Match 54.0%; Score 47; DB 18; Length 1007;
Best Local Similarity 53.8%; Pred. No. 69;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFWRL 15
|:|:|:|:|:
Db 7 YKSHLAYGAWAI 19

RESULT 5

US-10-450-763-56548
; Sequence 56548, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 56548
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(267)
; OTHER INFORMATION: xaa = X or * as defined in Table 2
US-10-450-763-56548

Query Match 50.6%; Score 44; DB 18; Length 267;
Best Local Similarity 53.8%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFW 13
|:|:|:|:|:
Db 120 RAYRADLRLEGW 132

RESULT 6

US-10-197-824-35
; Sequence 35, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
; FILE REFERENCE: 9U 106 R1
; CURRENT APPLICATION NUMBER: US/10/197,824
; CURRENT FILING DATE: 2002-07-19

; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 1356
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-197-824-35

Query Match 50.6%; Score 44; DB 15; Length 1356;
Best Local Similarity 58.3%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFW 14
|:|:|:|:|:
Db 280 YQHRRLYGQWK 291

RESULT 7

US-10-197-824-19
; Sequence 19, Application US/10197824
; Publication No. US20040023219A1
; GENERAL INFORMATION:
; APPLICANT: ORIGENE TECHNOLOGIES INC
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES
; FILE REFERENCE: 9U 106 R1
; CURRENT APPLICATION NUMBER: US/10/197,824
; CURRENT FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-197-824-19

Query Match 50.6%; Score 44; DB 15; Length 1593;
Best Local Similarity 58.3%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFW 14
|:|:|:|:|:
Db 517 YQHRRLYGQWK 528

RESULT 8

US-10-450-763-34690
; Sequence 34690, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 34690
; LENGTH: 1891
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (971)..(1000)
; OTHER INFORMATION: Kw REPLICATION REPB REPA INITIATION domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM00653A, p-value=7.207e-27, raw score
; OTHER INFORMATION: 20.40
US-10-450-763-34690

Query Match 50.6%; Score 44; DB 18; Length 1891;
Best Local Similarity 53.8%; Pred. No. 3.9e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFW 13

Db 120 RAYRADLRLEGYW 132

RESULT 9

```

US-11-028-539-34
; Sequence 34, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: RIGER, Jean-Pierre
; APPLICANT: RIEGER, François
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE
; OF ENDGENOUS RETROVIRA
; FILE REFERENCE: 200936USO/PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-028-539-34

```

Query Match 49.4%; Score 43; DB 19; Length 29;
Best Local Similarity 53.8%; Pred. No. 9.9;
Matches 7; Conservative 2; Mismatches 4; Indels

Qy	3	YQHD	LRA	YGF	WRL	15
		:	:	:	:	:
D _b	6	FSHD	PQAS	GLWRL	18	

RESULT 10

```

US-10-156-761-9242
; Sequence 9242, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10156761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9242
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9242

```

Query Match 49.4%; Score 43; DB 14; Length 256;
Best Local Similarity 72.7%; Pred. No. 82;
Matches 8; Conservative 1; Mismatches 2; Indels

Qy	2	SYQHDLRAYGF	12
		: : : : :	
Db	224	TYQHDLSPAGF	234

RESULT 11

```

US-11-028-539-36
; Sequence 36, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIGER, François
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL
; FILE REFERENCE: 200936USO/PCT
; CURRENT APPLICATION NUMBER: US/11/028,539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (85)..(85)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (165)..(165)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (187)..(187)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (330)..(330)
; OTHER INFORMATION: Xaa is any amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (360)..(360)
; OTHER INFORMATION: Xaa is any amino acid
US-11-028-539-36

```

Query Match	49.4%	Score 43;	DB 19;	Length 360;
Best Local Similarity	53.8%	Pred. No. 1.1e+02;		
Matches 7; Conservative	2;	Mismatches 4;	Indels	

Qy	3	YQHDLRAYGFWRL	15
		:	:
		:	:
Db	336	FSHDPQASGLWRL	348

RESULT 12

```

US-11-028-539-58
; Sequence 58, Application US/11028539
; Publication No. US20050118573A1
; GENERAL INFORMATION:
; APPLICANT: ALLIEL, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCING
; TITLE OF INVENTION: ENDOGENOUS RETROVIRAL ACTIVITY
; FILE REFERENCE: 200936USOPCT
; CURRENT APPLICATION NUMBER: US/11/028539
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: US/09/719179

```



```
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59950
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59950
```

```
Query Match          49.4%; Score 43; DB 15; Length 466;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      8 RYAGFWRL 15
Db      430 RRYGFWRL 437
```

```
RESULT 16
US-11-013-235-32
; Sequence 32, Application US/11013235
; Publication No. US20050136462A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenyu
; TITLE OF INVENTION: A Method for Engineering Nicking Enzymes
; FILE REFERENCE: NEB-235-US
; CURRENT APPLICATION NUMBER: US/11/013,235
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: US 60/531,064
; PRIOR FILING DATE: 2003-12-19
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 525
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: Bacillus stearothermophilus NUB36
US-11-013-235-32
```

```
Query Match          49.4%; Score 43; DB 20; Length 525;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy      1 RSYQHDRLRAYGFW 13
Db      225 RLYNHDRRAFWMW 237
```

```
RESULT 17
US-09-966-997-13
; Sequence 13, Application US/09966997
; Publication No. US20030100052A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Shuang-yong
; APPLICANT: Dore, Andrew
; APPLICANT: Hume, Adam
; APPLICANT: Pelletier, John
; APPLICANT: Zhou, Jing
; TITLE OF INVENTION: Method For Cloning And Expression Of BamBI Restriction
; TITLE OF INVENTION: Endonuclease And BamBI Methylase In E. coli And
; TITLE OF INVENTION: Purification Of BamBI Endonuclease
; FILE REFERENCE: NEB-191
; CURRENT APPLICATION NUMBER: US/09/966,997
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus B61
US-09-966-997-13
```

```
Query Match          49.4%; Score 43; DB 10; Length 530;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy      1 RSYQHDRLRAYGFW 13
Db      226 RLYNHDRRAFWMW 238
```

```
RESULT 18
US-10-425-114-59634
; Sequence 59634, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59634
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3606-053-G5_FLI.pep
US-10-425-114-59634
```

```
Query Match          48.9%; Score 42.5; DB 15; Length 413;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
```

```
Qy      2 SYQHDRLRAY---GFWR 14
Db      107 SYQKSLRAYRACGDWR 122
```

```
RESULT 19
US-10-425-115-194035
; Sequence 194035, Application US/10425115
; Publication No. US20040214272A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 194035
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_108540C.1.pep
US-10-425-115-194035

Query Match      48.9%; Score 42.5; DB 16; Length 461;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY      2 SYQHDLRAY---GFWR 14
      ||| |||| | |||
Db      155 SYQSLRAYRACGDWR 170

RESULT 20
US-10-425-115-233513
; Sequence 233513, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 233513
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144558C.1.pep
US-10-425-115-233513

Query Match      48.3%; Score 42; DB 16; Length 76;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 RSYQHDLRAYGF 12
      : ||| ||| |||
Db      19 KSYVHLRKYGF 30

RESULT 21
US-10-437-963-194505
; Sequence 194505, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194505
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90542C.1.pep
US-10-437-963-194505

Query Match      48.3%; Score 42; DB 16; Length 370;
Best Local Similarity 63.6%; Pred. No. 1.7e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 RSYQHDLRAYG 11
      : ||| ||| : ||
Db      182 RTYQHDLDLDFG 192

RESULT 22
US-10-767-701-42825
; Sequence 42825, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42825
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2047_1.pep
US-10-767-701-42825

Query Match      48.3%; Score 42; DB 16; Length 558;
Best Local Similarity 41.7%; Pred. No. 2.5e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      3 YQHDLRAYGFWR 14
      : ||| : |||
Db      167 FRHDINLYMFWR 178

RESULT 23
US-10-437-963-121169
; Sequence 121169, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
```

```

; SEQ ID NO 121169
;
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Oryza sativa
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24220C.1.pep
US-10-437-963-121169

Query Match      48.3%; Score 42; DB 16; Length 891;
Best Local Similarity 56.2%; Pred. No. 4e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Qy      6 DLRAYGF-----WRL 15
      |||:||||
Db      169 DLRSYGFFGILRLWRL 184

RESULT 24
US-10-972-963-136
; Sequence 136, Application US/10972963
; Publication No. US20050147993A1
; GENERAL INFORMATION:
; APPLICANT: Gencia Corporation
; APPLICANT: Khan, Shaharyar
; TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides
; FILE REFERENCE: 120701-2030
; CURRENT APPLICATION NUMBER: US/10/972,963
; CURRENT FILING DATE: 2004-10-25
; PRIOR APPLICATION NUMBER: 60/568,436
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: 60/513,983
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 136
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-972-963-136

Query Match      48.3%; Score 42; DB 18; Length 994;
Best Local Similarity 60.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      5 HDLRAYGFWR 14
      ||||:|
Db      403 HDLRDHAFWK 412

RESULT 25
US-10-425-115-246608
; Sequence 246608, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 246608
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_156490C.1.pep
US-10-425-115-246608

Query Match      47.1%; Score 41; DB 16; Length 63;

```



```
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 470
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: Bacillus stearothermophilus 6-55
US-11-013-235-37

Query Match          47.1%; Score 41; DB 20; Length 470;
Best Local Similarity 53.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 224 RSYTDRRAFEYW 236

RESULT 33
US-10-293-958-5
; Sequence 5, Application US/10293958
; Publication No. US20030140380A1
; GENERAL INFORMATION:
; APPLICANT: HENKES, STEFAN
; APPLICANT: CHEN, RUOYING
; APPLICANT: DA COSTA E SILVA, OSWALDO
; APPLICANT: VAN THIELEN, NOCHIA
; TITLE OF INVENTION: AMINE OXIDASE STRESS-RELATED POLYPEPTIDES AND METHODS
; TITLE OF INVENTION: OF USE IN PLANTS
; FILE REFERENCE: 16313-0174
; CURRENT APPLICATION NUMBER: US/10/293,958
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/345,969
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-293-958-5

Query Match          47.1%; Score 41; DB 14; Length 497;
Best Local Similarity 54.5%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFW 13
Db 115 YDHDLESYGLY 125

RESULT 34
US-11-013-235-31
; Sequence 31, Application US/11013235
; Publication No. US20050136462A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Shuangyong
; APPLICANT: Xu, Shuangyong
; TITLE OF INVENTION: A Method for Engineering Nicking Enzymes
; FILE REFERENCE: NEB-235-US
; CURRENT APPLICATION NUMBER: US/11/013,235
; CURRENT FILING DATE: 2004-12-15
; PRIOR FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: US 60/531,064
; PRIOR FILING DATE: 2003-12-19
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 524
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: Bacillus stearothermophilus 6-55
```

```
US-11-013-235-31

Query Match          47.1%; Score 41; DB 20; Length 524;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 225 RSYTDRRAFEYW 237

RESULT 35
US-10-106-275-6
; Sequence 6, Application US/10106275
; Publication No. US20030186363A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenyu
; APPLICANT: Xu, Shuangyong
; TITLE OF INVENTION: Method For Cloning And Expression of BsaI Restriction Endonuclease
; TITLE OF INVENTION: BsaI Methylase In E. Coli
; FILE REFERENCE: NEB-202
; CURRENT APPLICATION NUMBER: US/10/106,275
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus 6-55
US-10-106-275-6

Query Match          47.1%; Score 41; DB 14; Length 544;
Best Local Similarity 53.8%; Pred. No. 3.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 229 RSYTDRRAFEYW 241

RESULT 36
US-10-437-963-129380
; Sequence 129380, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 129380
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31641C.1.pep
US-10-437-963-129380

Query Match          47.1%; Score 41; DB 16; Length 573;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFWR 14
Db 111 : : : :
```

Db 263 RSYRGLSSGGWK 276

RESULT 37

US-10-220-335-578

Sequence 578, Application US/10220335

Publication No. US20050175607A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-041

CURRENT APPLICATION NUMBER: US/10/220,335

CURRENT FILING DATE: 2002-08-28

PRIOR APPLICATION NUMBER: 09/664,641

PRIOR FILING DATE: 2000-09-19

PRIOR APPLICATION NUMBER: 09/616,807

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 09/597,707

PRIOR FILING DATE: 2000-06-17

PRIOR APPLICATION NUMBER: 09/577,409

PRIOR FILING DATE: 2000-05-18

PRIOR APPLICATION NUMBER: 09/515,126

PRIOR FILING DATE: 2000-02-28

NUMBER OF SEQ ID NOS: 688

SOFTWARE: Custom

SEQ ID NO 578

LENGTH: 588

TYPE: PRT

ORGANISM: Homo sapiens

NAME/KEY: misc_feature

LOCATION: (1)...(588)

OTHER INFORMATION: Xaa = any amino acid or nothing

US-10-220-335-578

Query Match 47.1%; Score 41; DB 18; Length 588;

Best Local Similarity 53.8%; Pred. No. 3.9e+02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 YQHDRLRAYGFWR 15

Db 510 FTQTLKAYGDWRL 522

RESULT 38

US-10-450-763-38718

Sequence 38718, Application US/10450763

Publication No. US20050196754A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc

TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 790CIP3/US

CURRENT APPLICATION NUMBER: US/10/450,763

CURRENT FILING DATE: 2003-06-11

PRIOR APPLICATION NUMBER: PCT/US01/08631

PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 09/540,217

PRIOR FILING DATE: 2000-03-31

PRIOR APPLICATION NUMBER: 09/649,167

PRIOR FILING DATE: 2000-08-23

NUMBER OF SEQ ID NOS: 60736

SOFTWARE: Custom

SEQ ID NO 38718

LENGTH: 651

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: DOMAIN

LOCATION: (296)..(351)

OTHER INFORMATION: Heat shock hsp90 proteins family proteins domain identified

OTHER INFORMATION: by eMATRIX, accession number BL00298G, p-value=3.025e-31, raw sc

OTHER INFORMATION: of 24.57

FEATURE:

NAME/KEY: DOMAIN

LOCATION: (96)..(648)

OTHER INFORMATION: Hsp90 protein domain identified by Pfam, accession name

OTHER INFORMATION: HSP90, E-value=2.2e-160, Pfam score of 546.2

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(651)

OTHER INFORMATION: Xaa = X or * as defined in Table 2

US-10-450-763-38718

Query Match 47.1%; Score 41; DB 18; Length 651;

Best Local Similarity 53.8%; Pred. No. 4.3e+02;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 YQHDRLRAYGFWR 15

Db 574 FTQTLKAYGDWRL 586

RESULT 39

US-10-389-566-1468

Sequence 1468, Application US/10389566

Publication No. US20040025202A1

GENERAL INFORMATION:

APPLICANT: Monsanto Technology, LLC

APPLICANT: Laurie, Cathy C

TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants

FILE REFERENCE: 38-77(52900)D

CURRENT APPLICATION NUMBER: US/10/389,566

CURRENT FILING DATE: 2003-03-31

PRIOR APPLICATION NUMBER: US 60/365,301

PRIOR FILING DATE: 2002-03-15

PRIOR APPLICATION NUMBER: US 60/391,786

PRIOR FILING DATE: 2002-06-25

PRIOR APPLICATION NUMBER: US 60/392,018

PRIOR FILING DATE: 2002-06-26

NUMBER OF SEQ ID NOS: 2459

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1468

LENGTH: 692

TYPE: PRT

ORGANISM: Schizosaccharomyces pombe

US-10-389-566-1468

Query Match 47.1%; Score 41; DB 15; Length 692;

Best Local Similarity 57.1%; Pred. No. 4.5e+02;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 1 RSYQHDLRAYGFWR 14

Db 236 RSQRFRLTYPEWR 249

RESULT 40

US-10-389-566-1481

Sequence 1481, Application US/10389566

Publication No. US20040025202A1

GENERAL INFORMATION:

APPLICANT: Monsanto Technology, LLC

APPLICANT: Laurie, Cathy C

TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants

FILE REFERENCE: 38-77(52900)D

CURRENT APPLICATION NUMBER: US/10/389,566

CURRENT FILING DATE: 2003-03-31

PRIOR APPLICATION NUMBER: US 60/365,301

PRIOR FILING DATE: 2002-03-15

PRIOR APPLICATION NUMBER: US 60/391,786

PRIOR FILING DATE: 2002-06-25

PRIOR APPLICATION NUMBER: US 60/392,018

PRIOR FILING DATE: 2002-06-26

NUMBER OF SEQ ID NOS: 2459

SOFTWARE: PatentIn version 3.2

SEQ ID NO 1481

```
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-389-566-1481

Query Match      47.1%  Score 41; DB 15; Length 817;
Best Local Similarity 57.1%; Pred. No. 5.3e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 RSYQHDLRLAYGFWR 14
      ||| ||| |||
Db      362 RSFQFLRLTYPEWR 375
```

Search completed: September 16, 2005, 12:57:58
Job time : 168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 12:33:52 ; Search time 40 Seconds
(without alignment)

36.081 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDLRAYGFWRLL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR_79:*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	49.4	118	2	G81992
2	43	49.4	118	2	G81221
3	43	49.4	188	2	T02822
4	43	49.4	459	2	D81950
5	42	48.3	444	2	T24844
6	42	48.3	887	2	T03939
7	42	48.3	994	1	SNCLM
8	41	47.1	188	2	D82248
9	41	47.1	817	2	T43659
10	40	46.0	223	2	A84436
11	40	46.0	274	2	F96961
12	40	46.0	330	2	B96027
13	40	46.0	371	2	T21153
14	40	46.0	414	2	T19912
15	40	46.0	447	2	E72320
16	40	46.0	451	2	T15118
17	40	46.0	516	2	D96682
18	40	46.0	912	2	T31223
19	40	46.0	1177	2	AG0284
20	39.5	45.4	309	2	A95988
21	39.5	45.4	471	2	D95940
22	39	44.8	81	2	A75372
23	39	44.8	103	2	A53461
24	39	44.8	206	2	S08029
25	39	44.8	206	2	G86076
26	39	44.8	206	2	H91229
27	39	44.8	286	2	A55530
28	39	44.8	313	2	AR1973
29	39	44.8	330	2	A82586

30	39	44.8	330	2	B97368	hypothetical sugar
31	39	44.8	335	2	I36790	hypothetical prote
32	39	44.8	340	2	C83827	molybdopterin bios
33	39	44.8	433	1	B65058	fixC protein homol
34	39	44.8	502	2	B75287	sensor histidine k
35	39	44.8	505	2	F71266	probable tpr prote
36	39	44.8	514	2	H87465	sodium-galactoside
37	39	44.8	700	2	S09699	bib protein - frui
38	39	44.8	736	2	T06757	hypothetical prote
39	39	44.8	778	2	H81311	transcription-repa
40	39	44.8	1055	2	T31111	ATPase 1 (EC 3.6.1
41	39	44.8	1135	1	JQ1528	G2-G1 polyprotein
42	39	44.8	2108	2	S72458	sodium channel pro
43	38.5	44.3	159	2	AH2118	bacterioferritin c
44	38.5	44.3	307	2	F97639	hypothetical prote
45	38.5	44.3	307	2	AG2862	conserved hypotet
46	38.5	44.3	404	2	D82355	probable transposa
47	38	43.7	72	2	S11653	outer membrane pro
48	38	43.7	95	2	S45381	outer membrane pro
49	38	43.7	127	2	AH1900	hypothetical prote
50	38	43.7	151	2	AC0495	probable acetyltra
51	38	43.7	171	2	AF0601	outer membrane pro
52	38	43.7	171	2	F85590	outer membrane pro
53	38	43.7	171	2	D90740	outer membrane pro
54	38	43.7	171	2	I55173	outer membrane pro
55	38	43.7	190	2	PQ0536	nonstructural prot
56	38	43.7	212	2	A72334	conserved hypotet
57	38	43.7	224	2	F89789	cell division and
58	38	43.7	241	1	F69513	conserved hypotet
59	38	43.7	247	2	A72517	hypothetical prote
60	38	43.7	320	2	A40660	outer membrane pro
61	38	43.7	328	2	S24203	protein 1 - Neisae
62	38	43.7	352	2	G70335	heterodisulfide re
63	38	43.7	453	2	T22153	hypothetical prote
64	38	43.7	459	2	C81156	conserved hypotet
65	38	43.7	481	2	F96802	hypothetical prote
66	38	43.7	518	2	T22375	hypothetical prote
67	38	43.7	654	2	F83260	hypothetical prote
68	38	43.7	696	2	B87265	Kup system potassi
69	38	43.7	741	2	H81877	probable isocitrat
70	38	43.7	741	2	H81143	isocitrate dehydro
71	38	43.7	771	2	T21633	hypothetical prote
72	38	43.7	978	2	T16948	hypothetical prote
73	38	43.7	1107	2	S61667	probable membrane
74	38	43.7	1192	2	S69000	laminin gamma 2 ch
75	38	43.7	1193	2	A44018	laminin B2 chain
76	38	43.7	142	1	WMAD19	early E1B 17K prot
77	37	42.5	227	2	T16766	hypothetical prote
78	37	42.5	256	2	G70488	undecaprenol kinas
79	37	42.5	260	2	E87577	glutamine cyclotra
80	37	42.5	267	2	S52347	hypothetical prote
81	37	42.5	282	2	B82187	DNA ligase VC1542
82	37	42.5	284	2	B83236	conserved hypotet
83	37	42.5	306	2	D96708	unknown protein, 9
84	37	42.5	317	2	AD1542	oxidoreductases ho
85	37	42.5	317	2	AF1184	oxidoreductases ho
86	37	42.5	319	2	A95976	glucosyltransferas
87	37	42.5	319	2	S40175	ExoW protein - Rhi
88	37	42.5	327	2	A81839	probable periplasm
89	37	42.5	327	2	D81103	hypothetical prote
90	37	42.5	331	2	S76480	hypothetical prote
91	37	42.5	379	2	S42529	Opaque-2-related p
92	37	42.5	412	2	T14668	hypothetical prote
93	37	42.5	419	2	S56073	opaque-2 protein -
94	37	42.5	434	2	SS1644	secreted/adhesive
95	37	42.5	437	2	A34800	Opaque-2 protein -
96	37	42.5	460	2	S06022	regulatory protein
97	37	42.5	465	2	B82515	alpha-L-fucosidase
98	37	42.5	487	2	F86771	cardiolipin syntha
99	37	42.5	489	2	S68280	protein disulfide-
100	37	42.5	514	2	T06585	ammonium transport

ALIGNMENTS

RESULT 1	
G81992	
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain A NMA0019 [similarity] - Neisseria	
C:Species: Neisseria meningitidis	
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004	
C:Accession: G81992	
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel	
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel	
R:Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,	
Nature 404, 502-506, 2000	
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.	
A:Reference number: A81775; MUID:20222556; PMID:10761919	
A:Accession: G81992	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-118 <PAR>	
A:CROSS-references: UNIPROT:Q90X78; GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB8333	
A:Experimental source: serogroup A, strain Z2491	
C:Genetics:	
A:Gene: nuaA; NMA0019	
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3	
C:Keywords: membrane-associated complex; NAD; oxidoreductase	
Query Match 49.4%; Score 43; DB 2; Length 118;	
Best Local Similarity 70.0%; Pred. No. 5;	
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY 6 DLRAYGFWR 15	
DB 86 DLGAYGFWS 95	
RESULT 2	
G81221	
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain A NMB0241 [similarity] - Neisseria	
C:Species: Neisseria meningitidis	
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004	
C:Accession: G81221	
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.	
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;	
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Fizza, M.	
Science 287, 1809-1815, 2000	
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve	
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.	
A:Reference number: A81000; MUID:20175755; PMID:10710307	
A:Accession: G81221	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-118 <TET>	
A:CROSS-references: UNIPROT:Q9K1C3; GB:AB002381; GB:AB002098; NID:g7225455; PIDN:AAF4069	
A:Experimental source: serogroup B, strain MC58	
C:Genetics:	
A:Gene: NMB0241	
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3	
C:Keywords: membrane-associated complex; NAD; oxidoreductase	
Query Match 49.4%; Score 43; DB 2; Length 118;	
Best Local Similarity 70.0%; Pred. No. 5;	
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;	
QY 6 DLRAYGFWR 15	
DB 86 DLGAYGFWS 95	
RESULT 3	
T02822	
probable membrane protein L2759.14 [imported] - Leishmania major (strain Friedlin)	
C:Species: Leishmania major	
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004	
C:Accession: B81459; T02822	

R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.; F	
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999	
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c	
A:Reference number: A81455; MUID:99178987; PMID:10077609	
A:Accession: B81459	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-188 <PYL>	
A:CROSS-references: UNIPROT:Q25352; GB:AE001274; NID:g3264850; PIDN:AAC24644.1; PID:g140:	
A:Experimental source: strain MHOM/IL/81/Friedlin	
C:Genetics:	
A:Gene: L2759.14	
A:Map position: 1	
C:Keywords: transmembrane protein	
Query Match 49.4%; Score 43; DB 2; Length 188;	
Best Local Similarity 77.8%; Pred. No. 8.1;	
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY 5 HDLRAYGF 13	
DB 46 HDLMAIGFW 54	
RESULT 4	
DB1950	
probable transmembrane efflux protein NMA1022 [imported] - Neisseria meningitidis (strain	
C:Species: Neisseria meningitidis	
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004	
C:Accession: DB1950	
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel	
R:Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,	
Nature 404, 502-506, 2000	
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.	
A:Reference number: A81775; MUID:20222556; PMID:10761919	
A:Accession: DB1950	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-459 <PAR>	
A:CROSS-references: UNIPROT:Q9JV27; GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84291	
A:Experimental source: serogroup A, strain Z2491	
C:Genetics:	
A:Gene: NMA1022	
C:Superfamily: conserved hypothetical protein H11612	
Query Match 49.4%; Score 43; DB 2; Length 459;	
Best Local Similarity 50.0%; Pred. No. 21;	
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	
QY 2 SYQHDLRAYGF 13	
DB 416 AYRFDMGIYGF 427	
RESULT 5	
T24844	
hypothetical protein T11F9.1 - Caenorhabditis elegans	
C:Species: Caenorhabditis elegans	
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004	
C:Accession: T24844	
R:Lennard, N.	
submitted to the EMBL Data Library, June 1996	
A:Reference number: Z19941	
A:Accession: T24844	
A>Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: DNA	
A:Residues: 1-444 <WIL>	
A:CROSS-references: UNIPROT:Q22405; EMBL:Z74042; PIDN:CAA98536.1; GSPDB:GN00023; CESP:T1J	
A:Experimental source: clone T11F9	
C:Genetics:	
A:Gene: CESP:T11F9.1	
A:Map position: 5	
A:Introns: 19/1; 61/2; 130/3; 177/3; 347/3; 375/1; 413/3	

C;Keywords: phosphoric monoester hydrolase; trehalose biosynthesis

Query Match 47.1%; Score 41; DB 2; Length 817;
Best Local Similarity 57.1%; Pred. No. 81;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFWR 14

||| : ||| : ||| :
362 RSFQRLTPEWR 375

RESULT 10

A84436

probable glutathione S-transferase [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: A84436

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: A84436

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-223 <STO>

A;Cross-references: UNIPROT:Q92VQ4; GB:AE002093; NID:g3894170; PIDN:AAC78520.1; GSPDB:GN

C;Genetics:

A;Gene: At2g02380

A;Map position: 2

C;Superfamily: plaiace glutathione transferase

Query Match 46.0%; Score 40; DB 2; Length 223;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFWR 14

||| : ||| : ||| :
8 YQAKLKLYWR 19

RESULT 11

F96961

bacitracin resistance protein (baca) [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004

C;Accession: F96961

R;Melling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: F96961

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-274 <KUR>

A;Cross-references: UNIPROT:Q97LQ3; GB:AE001437; PIDN:AAK78481.1; PID:g15023362; GSPDB:G

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC0501

C;Superfamily: Escherichia coli bacitracin resistance protein baca

Query Match 46.0%; Score 40; DB 2; Length 274;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QHDLRAYGFWRL 15

||| : ||| : ||| :
250 RHDKFAFGWYRI 261

RESULT 12

B96027

probable aliphatic sulfonates uptake ABC transporter periplasmic solute-binding protein i

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C;Accession: B96027

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: B96027

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-330 <KUR>

A;Cross-references: UNIPROT:Q926E3; GB:AL591985; PIDN:CAC49882.1; PID:g15141370; GSPDB:G

A;Experimental source: strain 1021, megaplasmid pSymb

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMB20570

A;Genome: plasmid

Query Match 46.0%; Score 40; DB 2; Length 330;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFWR 14

||| : ||| : ||| :
179 RGNQVDIAAGYQW 192

RESULT 13

T21153

hypothetical protein F20E11.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T21153

R;Ainscough, R.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19383

A;Accession: T21153

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-371 <WIL>

A;Cross-references: UNIPROT:Q9XV86; EMBL:Z81508; PIDN:CAB04139.1; GSPDB:GN000023; CESP:F2

A;Experimental source: clone P20E11

C;Genetics:

A;Gene: CESP:F20E11.6

A;Map position: 5

A;Introns: 173/2; 211/3; 247/2

Query Match 46.0%; Score 40; DB 2; Length 371;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 RAYGFWR 15

||| : ||| : ||| :
70 RAYGFWR 77

RESULT 14

T19912

hypothetical protein C43F9.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T19912

R;Nortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19195
A:Accession: T19912
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-414 <WIL>
A:Cross-references: UNIPROT:Q9XUG7; EMBL:Z82262; PIDN:CAB05151.1; GSPDB:GN000022; CESP:C43P9
A:Experimental source: clone C43P9
C:Genetics:
A:Gene: CESP:C43P9.4
A:Map position: 4
A:Introns: 53/3; 81/2; 111/3; 134/3; 164/3; 219/1; 346/1

Query Match 46.0%; Score 40; DB 2; Length 414;

Best Local Similarity 60.0%; Pred. No. 59; Mismatches 1; Indels 3; Gaps 0;

QY 4 QHDLRAYGF 13
DB 21 KHSLMPYGF 30

RESULT 15

E72320
hypothetical protein TM0884 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
A:Accession: E72320
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <ARN>
A:Cross-references: UNIPROT:Q9WZY6; GB:AE001754; GB:AE000512; NID:G4981417; PIDN:AA03596
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0884

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

Query Match 46.0%; Score 40; DB 2; Length 447;
Best Local Similarity 66.7%; Pred. No. 63; Mismatches 3; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 5 HDLRAYGF 13
DB 11 HDPAYDF 19

RESULT 16

T15718
hypothetical protein C30G12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
A:Accession: T15718
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-451 <LAT>
A:Cross-references: EMBL:U21319; NID:G687832; PIDN:AAC46670.1; CESP:C30G12.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:C30G12.1
A:Introns: 37/3; 71/2; 107/3; 153/2; 205/1; 289/1; 376/3

Query Match 46.0%; Score 40; DB 2; Length 451;

Best Local Similarity 77.8%; Pred. No. 64; Mismatches 7; Conservative 1; Indels 1; Gaps 0;

QY 2 SYOHLRAY 10
DB 61 SYOHLUSAF 69

RESULT 17

D96682
protein F1E22.18 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
A:Accession: D96682

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96682

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-516 <STO>

A:Cross-references: UNIPROT:Q9SHX4; GB:AE005173; NID:G686400; PIDN:AAF23834.1; GSPDB:GN C:Genetics:
A:Gene: F1E22.18
A:Map position: 1

Query Match 46.0%; Score 40; DB 2; Length 516;

Best Local Similarity 54.5%; Pred. No. 74; Mismatches 2; Indels 3; Gaps 0;

QY 3 YQHLRAYGF 13
DB 115 YDHLSYGLF 125

RESULT 18

T31223
traG protein homolog - Sphingomonas aromaticivorans plasmid pNLI

C:Species: Sphingomonas aromaticivorans

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

A:Accession: T31223

R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G submitted to the EMBL Data Library, July 1998

A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati.

A:Reference number: Z20992

A:Accession: T31223

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-912 <ROM>

A:Cross-references: UNIPROT:O85931; EMBL:AF079317; NID:G3378261; PID:G3378364; PIDN:AA0 A:Genetics:
A:Gene: plasmid pNLI
A:Note: traG

Query Match 46.0%; Score 40; DB 2; Length 912;

Best Local Similarity 53.3%; Pred. No. 1.3e+02; Mismatches 8; Conservative 1; Indels 6; Gaps 0;

QY 1 RSYQHDLRAYGF 15
DB 537 RGYATDLRSQQWYL 551

RESULT 19

AG0284

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-Kb pSymB megaplasmid from the N2-fixing endo-

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: D95940

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-471 <CUR>

A;Cross-references: UNIPROT_Q92VC0; GB:A1591985; PIDN:CAC49188.1; PID:g15140673; GSPDB:G1

R;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

Lela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

P.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaullt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: Smb21284

A;Genome: plasmid

C;Keywords: oxidoreductase

	Query Match	45.4%	Score 39.5;	DB 2;	Length 471;
	Best Local Similarity	61.5%;	Pred. No. 81;		
	Matches	8;	Conservative	2;	Mismatches
				Indels	Gaps
Qy	3 QYHDLRAYGFWRLL 15	:	:	:	:
Dd	72 YEYGARA-GFWRL 83	:	:	:	:

RESULT 22

A75372

Hypothetical protein - Deinococcus radiodurans (strain RL)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: A75372

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F

S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.C.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RL.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: A75372

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-81 <WHI>

A;Cross-references: UNIPROT_Q9RTW0; GB:AE002007; GB:AE000513; NID:g6459402; PIDN:AAF11204

A;Experimental source: strain RL

C;Genetics:

A;Gene: DR1643

A;Map position: 1

	Query Match	44.8%	Score 39;	DB 2;	Length 81;
	Best Local Similarity	100.0%;	Pred. No. 16;		
	Matches	6;	Conservative	0;	Mismatches
				Indels	Gaps
Qy	9 AYGEFWR 14	:	:	:	:
Dd	64 AYGEFWR 69	:	:	:	:

RESULT 23

A53461

Voltage-sensitive sodium channel - house fly (fragment)

C;Species: Musca domestica (house fly)

C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C;Accession: A53461

R;Knipple, D.C.; Doyle, K.E.; Marsella-Herrick, P.A.; Soderlund, D.M.

F;Proc. Natl. Acad. Sci. U.S.A. 91, 2483-2487, 1994

A;Title: Tight genetic linkage between the kdr insecticide resistance trait and a voltage

A;Reference number: A53461; MUID:94195766; PMID:8146143

A;Contents: NAIDM, insecticide-susceptible
A;Accession: A53461
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <KNI>
A;Cross-references: UNIPROT:Q25439; UNIPROT:Q25440; UNIPROT:Q94615
A;Note: sequence inconsistent with nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:146081, NCBIP:146080)
C;Superfamily: sodium channel protein
C;Keywords: duplication

Query Match 44.8%; Score 39; DB 2; Length 103;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDLRAYGFWRLL 15
||| ||| |||
Db 46 YQHVLQAGPWHM 58

RESULT 24

hypothetical 23.5K protein (glnA-fdhE intergenic region) - Escherichia coli (strain K-12)
N;Alternate names: hypothetical protein o206
C;Species: Escherichia coli
C;Date: 06-Oct-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S40829; H65193
R;Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 21, 3391-3398, 1993
A;Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8
A;Reference number: S40802; MUID:93347969; PMID:8346018
A;Accession: S40829
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-206 <PLU>
A;Cross-references: UNIPROT:P32145; EMBL:L19201; NID:G304961; PIDN:AAB03018.1; PID:G5049
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H65193
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-206 <BLAT>
A;Cross-references: GB:AE000464; GB:U00096; NID:G2367324; PIDN:AAD13447.1; PID:G1790318;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yihX
A;Start codon: GTG
C;Superfamily: Alcaligenes eutrophus phosphoglycolate phosphatase

Query Match 44.8%; Score 39; DB 2; Length 206;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
||| ||| |||
Db 152 RIYQHVLAQEGF 163

RESULT 25

probable phosphatase yihX [Imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G86076
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G86076
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <STO>
A;Cross-references: UNIPROT:Q8X8D1; GB:AE005174; NID:G12518771; PIDN:AAG59075.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yihX
C;Superfamily: Alcaligenes eutrophus phosphoglycolate phosphatase

Query Match 44.8%; Score 39; DB 2; Length 206;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
||| ||| |||
Db 152 RIYQHVLAQEGF 163

RESULT 26

H91229
Probable phosphatase [Imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H91229
R;Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91229
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <HAY>
A;Cross-references: UNIPROT:Q8X8D1; GB:BA000007; PIDN:BA838231.1; PID:G13364284; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC84808
C;Superfamily: Alcaligenes eutrophus phosphoglycolate phosphatase

Query Match 44.8%; Score 39; DB 2; Length 206;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
||| ||| |||
Db 152 RIYQHVLAQEGF 163

RESULT 27

A55530
megakaryocyte growth and development factor, long form - human
N;Alternate names: MPl, ligand, long form
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C;Accession: A55530
R;Chang, M.; McNinch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs, S
J. Biol. Chem. 270, 511-514, 1995
A;Title: Cloning and characterization of the human megakaryocyte growth and development
A;Reference number: A55530; MUID:95122483; PMID:7822271
A;Accession: A55530
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-286 <CHA>
A;Cross-references: GB:U17071
C;Genetics:
A;Gene: MGDF
A;Map position: 3q26.3
C;Keywords: alternative splicing; cytokine

Query Match 44.8%; Score 39; DB 2; Length 286;
Best Local Similarity 58.3%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

Qy      2 SYQHDLRAYGFW 13
      1:|||||
Db      151 SFQHLRAGKDFW 162

RESULT 28
AF1973
hypothetical protein alr1337 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF1973
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF1973
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313 <KUR>
A:Cross-references: UNIPROT:Q8YX78; GB:BA000019; PIDN:BAB73294.1; PID:gl7130684; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1337
C:Superfamily: ADP-ribosylglycohydrolase

Query Match      44.8%; Score 39; DB 2; Length 313;
Best Local Similarity 50.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      1 RSYQHDLRAYGFW 14
      1:|||||
Db      69 RGYSLDAIHSFWR 82

RESULT 29
AB2586
pfkB family carbohydrate kinase [imported] - Agrobacterium tumefaciens (strain C58, Dup
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2586
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthrner, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2586
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <KUR>
A:Cross-references: UNIPROT:Q8UJ59; GB:AE008688; PIDN:AAL41104.1; PID:gl7738396; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0079
A:Map position: circular chromosome

Query Match      44.8%; Score 39; DB 2; Length 330;
Best Local Similarity 54.5%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      3 YQHDLRAYGFW 13
      1:|||||
Db      90 FQHDIRAQGVY 100

RESULT 30
B97368
hypothetical sugar kinase alr0537 [imported] - Agrobacterium tumefaciens (strain C58, Cd

```

```

C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97368
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97368
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <KUR>
A:Cross-references: UNIPROT:Q8U5P4; GB:AE007869; PIDN:AAK85899.1; PID:gl15154944; GSPDB:B
C:Genetics:
A:Gene: AGR_C118
A:Map position: circular chromosome

Query Match      44.8%; Score 39; DB 2; Length 330;
Best Local Similarity 54.5%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      3 YQHDLRAYGFW 13
      1:|||||
Db      90 FQHDIRAQGVY 100

RESULT 31
I36790
hypothetical protein ORF44 - ictaluriid herpesvirus 1 (strain auburn 1)
C:Species: ictaluriid herpesvirus 1
A:Note: host Ictalurus punctatus (channel catfish)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Jul-2004
C:Accession: I36790
R;Davison, A.J.
submitted to GenBank, January 1992
A:Description: Channel catfish virus: a new type of herpesvirus.
A:Reference number: A36804
A:Accession: I36790
A:Molecule type: DNA
A:Residues: 1-335 <DAV>
A:Cross-references: UNIPROT:Q00124; GB:M75136; NID:G331209; PIDN:AAA88147.1; PID:G331254
R;Davison, A.J.
Virology 186, 9-14, 1992
A:Title: Channel catfish virus: a new type of herpesvirus.
A:Reference number: A39447; MUID:92087490; PMID:1727613
A:Contents: annotation
A:Note: neither protein nor nucleic acid sequence is given
C:Genetics:
A:Gene: 44
C:Superfamily: ictaluriid herpesvirus 1 hypothetical protein ORF44

Query Match      44.8%; Score 39; DB 2; Length 335;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      5 HDLRAYGFWRL 15
      1:|||||
Db      249 YDLRCSCGWRL 259

RESULT 32
CB3827
molybdopterin biosynthesis moeB [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: CB3827
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: CB3827
A>Status: preliminary
A:Molecule type: DNA

```


Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2005, 12:25:16 ; Search time 167 Seconds

(without alignments)
45.995 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87
Sequence: 1 RSYQHLRLAYGFWRLL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : UniProt_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	54.0	533	Q8RNV4	Q8RNV4 hafnia alve
2	46	52.9	227	Q8T09	Q8T09 methanosaec
3	46	52.9	587	Q86HI9	Q86HI9 dictyosteli
4	45	51.7	374	Q68J85	Q68J85 escherichia
5	44	50.6	1056	Q64N19	Q64N19 bacteroides
6	44	50.6	1430	Q72Z88	Q72Z88 brachydanio
7	44	50.6	1478	1 THO2 HUMAN	Q8NI27 homo sapien
8	44	50.6	1504	Q6DRE1	Q6DRE1 brachydanio
9	43	49.4	118	Q9UX78	Q9UX78 neisseria m
10	43	49.4	118	Q9K1C3	Q9K1C3 neisseria m
11	43	49.4	188	Q25352	Q25352 leishmania
12	43	49.4	256	Q82MP7	Q82MP7 streptomyce
13	43	49.4	430	Q6J925	Q6J925 arthrolepti
14	43	49.4	459	1 NORM_NEIMA	Q6JY27 neisseria m
15	43	49.4	465	Q6UQ64	Q6UQ64 bacillus st
16	43	49.4	530	Q69B20	Q69B20 bacillus st
17	42	48.3	318	Q8TPC0	Q8TPC0 mycobacteri
18	42	48.3	321	Q67P30	Q67P30 symbiobacte
19	42	48.3	418	Q8TRX2	Q8TRX2 methanosaec
20	42	48.3	444	Q22405	Q22405 caenorhabdi
21	42	48.3	652	Q9VJH0	Q9VJH0 drosophila
22	42	48.3	880	1 SYA_BUCBP	P59420 buchnera ap
23	42	48.3	887	Q04242	Q04242 zea mays (m
24	42	48.3	907	Q8H569	Q8H569 oryza sativ
25	42	48.3	907	Q7MUS9	Q7MUS9 porphyromon
26	42	48.3	951	1 SYLM_NEUCR	P15181 neurospora
27	42	48.3	994	Q9NEG4	Q9NEG4 leishmania
28	41.5	47.7	306	Q721X3	Q721X3 botryllus s
29	41	47.1	60	Q721X3	Q721X3 homo sapien
30	41	47.1	61	Q9H356	Q9H356 homo sapien
31	41	47.1	61	Q8D683	Q8D683 vibrio vuln

32	41	47.1	188	2	Q9KT60	Q9KT60 vibrio chol
33	41	47.1	191	2	Q7MIS3	Q7MIS3 vibrio vuln
34	41	47.1	191	2	Q8DB49	Q8DB49 methanosaec
35	41	47.1	229	2	Q8PW04	Q8PW04 pseudomonas
36	41	47.1	260	2	Q8G977	Q8G977 pseudomonas
37	41	47.1	260	2	Q8GBV1	Q8GBV1 pseudomonas
38	41	47.1	301	1	RL5_NEUCR	RL5_NEUCR neurospora
39	41	47.1	323	2	Q6MAL4	Q6MAL4 paracitriamyd
40	41	47.1	338	2	Q925V9	Q925V9 pseudomonas
41	41	47.1	354	2	Q8ZVC3	Q8ZVC3 pseudomonas
42	41	47.1	368	2	Q73TGL	Q73TGL mycobacteri
43	41	47.1	404	2	Q52212	Q52212 pseudomonas
44	41	47.1	404	2	Q6XUM9	Q6XUM9 pseudomonas
45	41	47.1	404	2	Q7BV60	Q7BV60 incl/m plas
46	41	47.1	425	2	Q9X7J2	Q9X7J2 pseudomonas
47	41	47.1	474	2	Q7VF30	Q7VF30 helicobacte
48	41	47.1	490	2	Q6T6X7	Q6T6X7 pellona fla
49	41	47.1	491	2	Q6T6X8	Q6T6X8 engraulis j
50	41	47.1	501	2	Q75812	Q75812 ashbya goss
51	41	47.1	514	2	Q8IRP3	Q8IRP3 drosophila
52	41	47.1	544	2	Q6SPF4	Q6SPF4 bacillus st
53	41	47.1	692	2	Q8TGH5	Q8TGH5 schizosacch
54	41	47.1	718	2	Q6FTA3	Q6FTA3 candida gla
55	41	47.1	817	1	TPPI3_SCHPO	TPPI3_SCHPO schizosacch
56	41	47.1	854	2	Q9FVG6	Q9FVG6 zea mays (m
57	41	47.1	944	2	Q6N0P0	Q6N0P0 rhodospseudo
58	41	47.1	150	2	Q82CE8	Q82CE8 streptomyce
59	40	46.0	191	2	Q87NM1	Q87NM1 vibrio para
60	40	46.0	200	2	Q89XB1	Q89XB1 bradyrhizob
61	40	46.0	206	2	Q72EP8	Q72EP8 leptospira
62	40	46.0	206	2	Q8F6X3	Q8F6X3 leptospira
63	40	46.0	215	2	Q6SU36	Q6SU36 mannheimia
64	40	46.0	223	1	GT22_ARATH	GT22_ARATH streptomyce
65	40	46.0	251	2	Q9LIN0	Q9LIN0 streptomyce
66	40	46.0	274	1	UPP1_CLOAB	UPP1_CLOAB clostridium
67	40	46.0	296	2	Q7N188	Q7N188 gloeobacter
68	40	46.0	310	2	Q6DOW7	Q6DOW7 erwinia car
69	40	46.0	323	2	Q90XK7	Q90XK7 lepiosteus
70	40	46.0	330	2	Q926E3	Q926E3 rhizobium m
71	40	46.0	361	2	Q9XUG7	Q9XUG7 caenorhabdi
72	40	46.0	371	2	Q9XV86	Q9XV86 caenorhabdi
73	40	46.0	438	2	Q7NHP5	Q7NHP5 gloeobacter
74	40	46.0	447	2	Q9XEA7	Q9XEA7 pinus radia
75	40	46.0	469	1	YQB1_CABEL	YQB1_CABEL thermotoga
76	40	46.0	497	2	Q94JZ7	Q94JZ7 caenorhabdi
77	40	46.0	497	2	Q8H191	Q8H191 arabidopsis
78	40	46.0	516	2	Q9SHX4	Q9SHX4 arabidopsis
79	40	46.0	518	2	Q9F915	Q9F915 arabidopsis
80	40	46.0	535	2	Q72RM5	Q72RM5 treponema p
81	40	46.0	535	2	Q8F439	Q8F439 leptospira
82	40	46.0	619	2	Q9C7B1	Q9C7B1 arabidopsis
83	40	46.0	637	2	Q9LHB0	Q9LHB0 arabidopsis
84	40	46.0	637	2	Q8CBD5	Q8CBD5 mus musculu
85	40	46.0	655	2	Q7QF16	Q7QF16 anopheles g
86	40	46.0	716	2	Q6CN86	Q6CN86 kluyveromyc
87	40	46.0	809	2	Q8QNB5	Q8QNB5 ectocarpus
88	40	46.0	863	2	Q62WC4	Q62WC4 bacillus li
89	40	46.0	886	2	Q65KX7	Q65KX7 bacillus li
90	40	46.0	912	2	Q85931	Q85931 sphingomona
91	40	46.0	1177	2	Q66A79	Q66A79 yersinia ps
92	40	46.0	1177	2	Q8ZE50	Q8ZE50 yersinia pe
93	40	46.0	1177	2	Q8ZAN1	Q8ZAN1 xylella fas
94	40	46.0	3377	2	Q8RTM4	Q8RTM4 homo sapien
95	40	46.0	8407	2	Q7RTM4	Q7RTM4 homo sapien
96	39.5	45.4	8797	1	SNE1_HUMAN	SNE1_HUMAN rhizobium m
97	39.5	45.4	309	2	Q92UG2	Q92UG2 rhizobium m
98	39.5	45.4	317	2	Q8GHM0	Q8GHM0 pseudomonas
99	39.5	45.4	471	2	Q92VC0	Q92VC0 rhizobium m
100	39.5	45.4	472	2	Q98C18	Q98C18 rhizobium l

ALIGNMENTS

```

RESULT 1
ID Q8RNY4 PRELIMINARY; PRT; 533 AA.
AC Q8RNY4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Restriction endonuclease.
GN Name=esp31R;
OS Hafnia alvei;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Hafnia.
OX NCBI_TaxID=569;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REFL3;
RX MEDLINE=22162868; PubMed=12172806; DOI=10.1007/s00438-002-0701-6;
RA Bitinaite J., Mitkaite G., Dauksaite V., Jakubauskas A., Timinskas A.,
RA Valsvila R., Lubys A., Janulaitis A.;
RT "Evolutionary relationship of Alw26I, Eco31I and Esp31I, restriction
RT endonucleases that recognise overlapping sequences.";
RL Mol. Genet. Genomics 267:664-672(2002).
DR EMBL: AF458983; AAM09641.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
KW Endonuclease.
SQ SEQUENCE 533 AA; 61708 MW; 5691332C65A6AD60 CRC64;

Query Match 54.0%; Score 47; DB 2; Length 533;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 225 RSYSHDRRSPEWM 237

RESULT 2
ID Q8TT09 PRELIMINARY; PRT; 227 AA.
AC Q8TT09;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein MA0634.
GN OrderedLocuNames=NA0634;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Athoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL: AE010725; AAM04076.1; -.
DR Pfam; PF02592; DUF165; 1.
DR TIGRFAMs; TIGR00697; DUF165; 1.
KW Complete proteome.
SQ SEQUENCE 227 AA; 25462 MW; 39E2EB615EB7CAB6 CRC64;

Query Match 52.9%; Score 46; DB 2; Length 227;
Best Local Similarity 51.5%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGFWR 14
Db 120 SYQHDRLWAFNWR 132

RESULT 3
ID Q86HI9 PRELIMINARY; PRT; 587 AA.
AC Q86HI9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to Dictyostelium discoideum (Slime mold). Phosphatidylinositol
DE 3-kinase 3 (EC 2.7.1.137) (PI3-kinase) (PtdIns-3-kinase) (PI3K).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.P., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117072; AAO52597.2; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016303; F:phosphatidylinositol 3-kinase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Kinase.
SQ SEQUENCE 587 AA; 66683 MW; 4B960B35F9379E06 CRC64;

Query Match 52.9%; Score 46; DB 2; Length 587;
Best Local Similarity 53.3%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFWR 15
Db 289 RSGRHLRCNGYQL 303

RESULT 4
ID Q68J85 PRELIMINARY; PRT; 374 AA.
AC Q68J85;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE TrbB.
GN Name=trbB;
OS Escherichia coli.
OG Flammid pSERB1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Dudley E.G., Abe C., Ghigo J.-M., Nataro J.P.;
RT "Type IV pilus mediated adherence of the atypical enteroaggregative
RT Escherichia coli strain C1096 to cultured cells and abiotic
RT surfaces.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY686591; AAT94232.1; -.
DR InterPro; IPR001854; Ribosomal_L29.
DR PROSITE; PS00579; RIBOSOMAL_L29; UNKNOWN_1.
KW Plasmid.
SQ SEQUENCE 374 AA; 41006 MW; ABAEB985A412B121 CRC64;

Query Match 51.7%; Score 45; DB 2; Length 374;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 QHDLRAYGFWR 14
Db 96 KHPSRKYGFWR 106

RESULT 5
Q64N19
ID Q64N19 PRELIMINARY; PRT; 1056 AA.
AC Q64N19;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Putative outer membrane protein probably involved in nutrient
DE binding.
GN ORFNames=BF4380;
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
[1]
RN SEQUENCE FROM N.A.
RP STRAINS=YCH46;
RC Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
RA Kuhara S., Hattori M., Hayashi T., Ohnishi Y.;
RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
RT inversions regulating cell surface adaptation.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0 (2004).
DR EMBL; AP006841; BAD51118.1; -.
SQ SEQUENCE 1056 AA; 117353 MW; F1B335D1D41B988A CRC64;

Query Match 50.6%; Score 44; DB 2; Length 1056;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SYQHDRLAYGFW 13
Db 388 SYEHNIRSGAFW 399

RESULT 6
Q72Z88
ID Q72Z88 PRELIMINARY; PRT; 1430 AA.
AC Q72Z88;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SI:B21018.2 (Novel protein similar to human THO complex 2 (THOC2))
DE (Fragment).
GN Name=thoc2; Synonyms=SI:B21018.2;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RP Babbage A.;

```

```

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL627383; CAD60810.1; -.
DR ZFIN; ZDB-GENE-030616-54; thoc2.
FT NON_TER 1 1
SQ SEQUENCE 1430 AA; 163500 MW; A31ADC92D23216AB CRC64;

Query Match 50.6%; Score 44; DB 2; Length 1430;
Best Local Similarity 58.3%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLAYGFWR 14
Db 424 YQHRYLQGWK 435

RESULT 7
THO2_HUMAN
ID THO2_HUMAN STANDARD; PRT; 1478 AA.
AC Q8N127; Q9H816;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE THO complex subunit 2 (Tho2).
GN Name=THOC2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND INTERACTION WITH THE
RP TREX COMPLEX.
RX MEDLINE=22010388; PubMed=1179277; DOI=10.1038/417304a;
RA Straesser K., Masuda S., Mason P., Pfannstiel J., Oppizzi M.,
RA Rodriguez-Navarro S., Rondon A.G., Aguilera A., Struhl K., Reed R.,
RA Hurt E.;
RT "TREX is a conserved complex coupling transcription with messenger RNA
RT export.";
RN Nature 417:304-308 (2002).
[2]
RP SEQUENCE OF 444-1162 FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., E.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikeda Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Yugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RN Nat. Genet. 36:40-45 (2004).

```

```
[3]
RN  SEQUENCE OF 1151-1478 FROM N.A. (ISOFORM 2).
RP  TISSUE=Salivary gland;
RA  Ottenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.-W., Weil B.,
RA  Anid C., Osanger A., Fobo G., Han M., Wiemann S.;
RA  Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RL  SubMITTED (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: The THO/TRex complex is recruited to transcribed genes
CC  and travels with the RNA polymerase during elongation. It may
CC  physically link proteins that function in transcription and in RNA
CC  export.
CC  -1- SUBUNIT: Part of the heteromultimeric THO/TRex complex containing
CC  THO1, THO2, THO3, THO4 and NFX1/UP56.
CC  -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC  -1- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=2;
CC  Name=1;
CC  IsoId=Q8NI27-1; Sequence=Displayed;
CC  Name=2;
CC  IsoId=Q8NI27-2; Sequence=VSP_008587, VSP_008588;
CC  Note=No experimental confirmation available;
CC  -1- SIMILARITY: Belongs to the THO2 family.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AF441770; AM28436.1; -.
DR  EMBL; AK023659; BAB14630.1; ALT_INIT.
DR  EMBL; BX648654; CAB46196.1; -.
DR  Genew; HGNC:19073; THO2.
DR  MIM; 300395; -.
DR  InterPro; IPR011001; Saposin like.
DR  TIGRfams; TIGR01847; bacteriocin_ssg; 1.
KW  Alternative splicing; mRNA processing; mRNA splicing; mRNA transport;
KW  Nuclear protein; RNA-binding; Transport.
FT  DOMAIN 1149 1476 Lys-rich.
FT  DOMAIN 808 813 Nuclear localization signal (Potential).
FT  VARSPLIC 1311 1314 /FTId=VSP_008587.
FT  VARSPLIC 1315 1478 Missing (in isoform 2).
FT  CONFLICT 692 692 F -> S (in Ref. 2).
FT  CONFLICT 1161 1161 E -> G (in Ref. 3).
FT  CONFLICT 1241 1241 E -> K (in Ref. 3).
SQ  SEQUENCE 1478 AA; 169580 MW; FEICBCE28C560BCF CRC64;
Query Match 50.6%; Score 44; DB 1; Length 1478;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 3 YQHLRAYGFWR 14
DB 402 YQHYRLYGQWK 413
RESULT 8
ID Q6DRE1 PRELIMINARY; PRT; 1504 AA.
AC Q6DRE1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tho2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
```

```
RP  SEQUENCE FROM N.A.
RX  PubMed=15256591; DOI=10.1073/pnas.0403929101;
RA  Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,
RA  Hopkins N.;
RT  "Identification of 315 genes essential for early zebrafish
RT  development.";
RL  Proc. Natl. Acad. Sci. U.S.A. 101:12792-12797(2004).
DR  EMBL; AY648818; AAT68136.1; -.
SQ  SEQUENCE 1504 AA; 172235 MW; B07A447C6B69EA99 CRC64;
Query Match 50.6%; Score 44; DB 2; Length 1504;
Best Local Similarity 58.3%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 3 YQHLRAYGFWR 14
DB 498 YQHYRLYGQWK 509
RESULT 9
ID Q9JX78 PRELIMINARY; PRT; 118 AA.
AC Q9JX78;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase I chain A (EC 1.6.5.3).
GN Name=nuoA; OrderedLocusNames=NNA0019;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=22491 / Serogroup A / Serotype 4A;
RX  MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA  Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA  Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA  Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin M., Holooyd S.,
RA  Jagals K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA  Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA  Whitehead S., Spratt B.G., Barrell B.G.;
RT  "Complete DNA sequence of a serogroup A strain of Neisseria
RT  meningitidis 22491.";
RL  Nature 404:502-506(2000).
CC  -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC  sulfur (Fe-S) centers, to quinones in the respiratory chain (By
CC  similarity).
CC  -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC  -1- SIMILARITY: Belongs to the complex I subunit 3 family.
DR  EMBL; AL162752; CAB83339.1; -.
DR  PIR; G81992; G81992.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR  GO; GO:0016491; F:oxidoreductase activity; IEA.
DR  GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR  InterPro; IPR00440; Oxidored_q4.
DR  Pfam; PF0507; Oxidored_q4; 1.
KW  Complete proteome; NAD; Oxidoreductase; Quinone; Transmembrane.
SQ  SEQUENCE 118 AA; 13584 MW; 38C401130EDDAE17 CRC64;
Query Match 49.4%; Score 43; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 6 DLRAYGFWR 15
DB 86 DLGAYGFWSM 95
RESULT 10
Q9K1C3
```

Q9KLC3 PRELIMINARY; PRT; 118 AA.
AC Q9KLC3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase I, A subunit.
GN OrderedLocNames=NM0241;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA Tetelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Citton H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizzo M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-sulfur (Fe-S) centers, to quinones in the respiratory chain (By similarity).
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the complex I subunit 3 family.
DR EMBL; AB002381; AAF040695.1; -;
DR PIR; G81221; G81221.
DR TIGR; NMB0241; -;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR00440; Oxidored_q4.
DR Pfam; PF00507; Oxidored q4; 1.
KW Complete proteome; NAD; Oxidoreductase; Quinone; Transmembrane.
SQ SEQUENCE 118 AA; 13598 MW; 21F58B3BAC579F84 CRC64;
Query Match 49.4%; Score 43; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 6 DIRAYGFWRLL 15
DB 86 DLGAYGFWSM 95
RESULT 11
ID Q25352 PRELIMINARY; PRT; 188 AA.
AC Q25352; O02505;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE L2759.14.
GN Name=L2759.14;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=99178987; PubMed=10077609; DOI=10.1073/pnas.96.6.2902;
RA Myler P.J., Audleman L., deVos T., Hixson G., Kiser P., Lemley C.,
RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;

RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-coding genes.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AE001274; AAC24644.1; -;
DR PIR; B81459; T02822.
DR InterPro; IPR008506; DUF788.
DR Pfam; PF05620; DUF788; 1.
SQ SEQUENCE 188 AA; 20964 MW; EE4A9F2B776E093B CRC64;
Query Match 49.4%; Score 43; DB 2; Length 188;
Best Local Similarity 77.8%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 HDLRAYGF 13
DB 46 HDLMAIGFW 54
RESULT 12
ID Q82MF7 PRELIMINARY; PRT; 256 AA.
AC Q82MF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative protocatechuate 3,4-dioxygenase beta subunit.
GN Name=pcdh; OrderedLocNames=SAV1703;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces avermitilis; deducing the ability of producing secondary metabolites.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005027; BAC69414.1; -;
DR HSSP; P00437; 3PCC.
DR GO; GO:0003824; F:Catalytic activity; IEA.
DR GO; GO:0008199; F:ferric iron binding; IEA.
DR GO; GO:0016702; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000627; Dioxxygenase.
DR Pfam; PF00775; Dioxxygenase C; 1.
DR PROSITE; PS00083; INTRADIOL_DIOXYGENAS; UNKNOWN_1.
KW Complete proteome; Dioxxygenase.
SQ SEQUENCE 256 AA; 28537 MW; 73592DB9F6DAC8ED CRC64;
Query Match 49.4%; Score 43; DB 2; Length 256;
Best Local Similarity 72.7%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 SYQHDLRAYGF 12
DB 224 TYQHLSAPGF 234
RESULT 13

```
Q6J925 Q6J925 PRELIMINARY; PRT; 430 AA.
AC Q6J925;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Recombination activating protein 1 (Fragment).
GN Name=RAG-1;
OS Arthroleptis variabilis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Microhylidae;
OC Arthroleptidae; Arthroleptis.
OX NCBI_TaxID=83977;
RN [1]
RP SEQUENCE FROM N.A.
RA van der Meijden A., Vences M., Meyer A.;
RT "Novel phylogenetic relationships of the enigmatic brevicipitine and
RT scaphiophryne toads as revealed by sequences from the nuclear Rag-1
RT gene.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2004).
DR EMBL; AY571642; AAT29041.1; -.
DR InterPro; IPR010988; M_synth_B12_like.
DR NON_TER 1
FT NON_TER 430
FT NON_TER 430
SQ SEQUENCE 430 AA; 49116 MW; 993256C18AF51AAE CRC64;

Query Match 49.4%; Score 43; DB 2; Length 430;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFW 14
||: ||: ||: ||:
DB 234 RSHENLQRYGTWR 247

RESULT 14
ID NORM_NEIMA STANDARD; PRT; 459 AA.
AC Q9JVT7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable multidrug resistance protein norm (Na(+)/drug antiporter)
DE (Multidrug-efflux transporter).
GN Names=norm; OrderedLocusNames=NMA1022;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devile K., Feltwell T., Hamlin N., Holooyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491";
RL Nature 404:502-506(2000).
CC -1- FUNCTION: Functions as a Na(+)/drug antiporter (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: Belongs to the multi antimicrobial extrusion (MATE)
CC family. Norm (TC 2.A.66.1.1) subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AL162754; CAB84291.1; -.
DR PIR; D81950; D81950.
DR HAMAP; MF 00400; -. 1.
DR InterPro; IPR002528; MatE.
DR Pfam; PF01554; MatE; 2.
DR TIGRFAMs; TIGR00797; matE; 1.
KW Antiport; Complete proteome; Inner membrane; Sodium transport;
KW Transmembrane; Transport.
FT TRANSMEM 20 40
FT TRANSMEM 53 73
FT TRANSMEM 100 120
FT TRANSMEM 132 152
FT TRANSMEM 168 188
FT TRANSMEM 202 222
FT TRANSMEM 252 272
FT TRANSMEM 285 305
FT TRANSMEM 325 345
FT TRANSMEM 358 378
FT TRANSMEM 395 415
FT TRANSMEM 423 443
SQ SEQUENCE 459 AA; 49941 MW; B2EFD115C8EA494E CRC64;

Query Match 49.4%; Score 43; DB 1; Length 459;
Best Local Similarity 50.0%; Pred. No. 83;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 13
||: ||: ||: ||:
DB 416 AYRFDMGYGF 427

RESULT 15
ID Q6U064 PRELIMINARY; PRT; 465 AA.
AC Q6U064;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BsmAI endonuclease.
GN Name=bmAIR;
OS Bacillus stearothermophilus
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhu Z., Zhou J., Xu S.-Y.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY364324; AAQ72363.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
KW Endonuclease.
SQ SEQUENCE 465 AA; 54697 MW; 3B72FD0F9AB3552C CRC64;

Query Match 49.4%; Score 43; DB 2; Length 465;
Best Local Similarity 53.8%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFW 13
||: ||: ||: ||:
DB 214 KSYTKDRRAYEY 226

RESULT 16
ID Q69B20 PRELIMINARY; PRT; 530 AA.
AC Q69B20;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE BsmBI restriction endonuclease.
GN Name=bsmBIR;
```

```

OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
ON NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B61;
RA Xu S.-Y., Dore A., Hume A., Pelletier J., Zhou J.;
RT "Method for cloning and expression of BmbI restriction endonuclease
RT and BmbI methylase in E. coli and purification of BmbI
RT endonuclease.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY452068; AAS0914.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
KW Endonuclease.
SQ SEQUENCE 530 AA; 62199 MW; 1443B3954B884B77 CRC64;

Query Match 49.4%; Score 43; DB 2; Length 530;
Best Local Similarity 53.8%; Pred. No. 97;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy . 1 RSYQHDLRAYGF 13
| : | | | | | : |
Db 226 RLYNHDRRAFMW 238

RESULT 17
ID Q8TPC0 PRELIMINARY; PRT; 318 AA.
AC Q8TPC0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Predicted protein.
GN OrderedLocusNames=MA1993;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobria; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
ON NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nuebaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Scange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010884; AAM05396.1; -.
KW Complete proteome.
SQ SEQUENCE 318 AA; 36371 MW; 377FBCF08D6576E0 CRC64;

Query Match 48.3%; Score 42; DB 2; Length 318;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 HDLRAYGF 13
| : | | | |
Db 223 HELISYGF 231

RESULT 18
ID Q742P0 PRELIMINARY; PRT; 318 AA.
AC Q742P0;

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP0795;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
ON NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ki0;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017229; AAS03112.1; -.
DR InterPro; IPR011251; Luciferase like.
DR Pfam; PF00236; Bac_luciferase; 1.
KW Complete proteome.
SQ SEQUENCE 318 AA; 34907 MW; 867FE56C1817A5AF CRC64;

Query Match 48.3%; Score 42; DB 2; Length 318;
Best Local Similarity 60.0%; Pred. No. 85;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YQHDLRAYGF 12
| : | | | |
Db 106 YEHEWYAYG 115

RESULT 19
ID Q67P30 PRELIMINARY; PRT; 321 AA.
AC Q67P30;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=STH1578;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
ON NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Iehikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium
RT thermophilum.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006840; BAD40563.1; -.
KW Hypothetical protein.
SQ SEQUENCE 321 AA; 35228 MW; D974E5E125135E2B CRC64;

Query Match 48.3%; Score 42; DB 2; Length 321;
Best Local Similarity 63.6%; Pred. No. 86;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 QHDLRAYGF 14
| : | | | |
Db 200 QHGRQAYQFWR 210

RESULT 20
ID Q8TRX2 PRELIMINARY; PRT; 418 AA.
AC Q8TRX2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Predicted protein.
GN OrderedLocusNames=MA1043;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobria; Methanosarcinales;

```

Query Match 48.3%; Score 42; DB 2; Length 444;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 6; Gaps

QY 1 RSYQHLRA-----YGFWRLL 15
|||:|:|:|:|:
DB 189 REVQLNSALARYEIVGYWRV 209
|||:|:|:|:|:
PRELIMINARY; PRT; 652 AA.

RESULT 22

Q9VJ0H
ID Q9VJ0H PRELIMINARY; PRT; 652 AA.
AC Q9VJ0H;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG13283-PA.
GN ORFNames=CG13283;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RX [1]
SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards M.G., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.H., Blazer J.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Arkil J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gong P., Gorrell J.G., Gu Z., Guan P., Harris M.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Scapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.D., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RP [2]
SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celnikier S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards M., Sodergren E.J.,
RA Swirskas R., Taber P.E., Wan K., Scapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

DR InterPro; IPR003938; EAG_ELK_ERG.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR001622; K+channel_pore.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00023; Ank; 5.
 DR Pfam; PF00027; CNMP binding; 1.
 DR Pfam; PF00520; Ion_Crns; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR PRINTS; PR01463; EAGCHANLFMLY.
 DR SMART; SM00248; ANK; 5.
 DR SMART; SM00100; CNMP; 1.
 DR PROSITE; PSS0088; ANK_REPEAT; 2.
 DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PSS0042; CNMP_BINDING; 3; 1.
 KW ANK repeat; Ion transport; Ionic channel; Transmembrane; Transport.
 SQ SEQUENCE 887 AA; 99265 MW; D8560520A6B419D CRC64;

Query Match 48.3%; Score 42; DB 2; Length 887;
 Best Local Similarity 56.2%; Pred. No. 2.5e+02;
 Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

QY 6 DLRAYGF-----WRL 15
 |||:||||
 Db 163 DLRSYGFFNMLRLWRL 178

RESULT 25
 Q8H569 PRELIMINARY; PRT; 907 AA.
 AC Q8H569;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative AKT1-like potassium channel.
 GN Name=OJ1656_E11.135;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Saaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 clone:OJ1656_E11."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003843; BAC24865.1; -.
 DR HSSP; Q60778; IOY3.
 DR Gramene; Q8H569; -.

DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005249; F:voltage-gated potassium channel activity; IEA.
 DR GO; GO:0006812; P:cation transport; IEA.
 DR GO; GO:0006813; P:potassium ion transport; IEA.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000595; CNMP binding.
 DR InterPro; IPR003938; EAG_ELK_ERG.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR001622; K+channel_pore.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00023; Ank; 5.
 DR Pfam; PF00027; CNMP binding; 1.
 DR Pfam; PF00520; Ion_Crns; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR PRINTS; PR01463; EAGCHANLFMLY.
 DR SMART; SM00248; ANK; 4.
 DR SMART; SM00100; CNMP; 1.
 DR PROSITE; PSS0088; ANK_REPEAT; 3.
 DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PSS0042; CNMP_BINDING; 3; 1.
 KW ANK repeat; Ion transport; Ionic channel; Transmembrane; Transport.
 SQ SEQUENCE 907 AA; 99461 MW; A0D9BAB721BD6E6E CRC64;

Query Match 48.3%; Score 42; DB 2; Length 907;
 Best Local Similarity 56.2%; Pred. No. 2.5e+02;
 Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Best Local Similarity 56.2%; Pred. No. 2.5e+02;
 Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

QY 6 DLRAYGF-----WRL 15
 |||:||||
 Db 169 DLRSYGFFGILRLWRL 184

RESULT 26
 Q7MUS9 PRELIMINARY; PRT; 951 AA.
 AC Q7MUS9;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PG1405;
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=W83;
 RX MEDLINE=22829867; PubMed=12949112;
 RX DOI=10.1128/JB.185.18.5591-5601.2003;
 RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
 RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
 RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
 RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
 RA Dewhirst F.E., Fraser C.M.;
 RT "Complete genome sequence of the oral pathogenic bacterium
 Porphyromonas gingivalis strain W83."
 RL J. Bacteriol. 185:5591-5601(2003).
 DR EMBL; AB017176; AAQ66463.1; -.
 DR TIGR; PG1405; -.

KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 951 AA; 108033 MW; D359CAB1A51C3C7B CRC64;

Query Match 48.3%; Score 42; DB 2; Length 951;
 Best Local Similarity 56.2%; Pred. No. 2.6e+02;
 Matches 9; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 2 SYQHDL--RAYGFWRLL 15
 |||:||||
 Db 599 SYMPDFTKRRYGFWEEL 614

RESULT 27
 SYLM_NEUCR STANDARD; PRT; 994 AA.
 AC P15181; Q7RVC8;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucyl-tRNA synthetase, mitochondrial precursor (EC 6.1.1.4)
 DE (leucine--tRNA ligase) (leuRS).
 GN Name=leu-5; ORFNames=99H12.170, NCU03814.1;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=90097874; PubMed=2574823;
 RA Chow C.M., Metzberg R.L., Rajbandary U.L.;
 RT "Nuclear gene for mitochondrial leucyl-tRNA synthetase of Neurospora
 crassa: isolation, sequence, chromosomal mapping, and evidence that
 the leu-5 locus specifies structural information."
 RL Mol. Cell. Biol. 9:4631-4644(1989).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=74-OR23-1A / FGSC 987;

```

RX MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gkg293;
RA Mannhaupt G., Montrone C., Haase D., Mewes H.-W., Aign V.,
RA Hohnsbeil J.D., Fartmann B., Nyakatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence."
RL Nucleic Acids Res. 31:1944-1954(2003).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich S., Smirnov S., Purcell S., Rehman B.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Butler J., Endrizzi M.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Iankiev P., Bell-Federsen D., Nelson M.A.,
RA Werner-Washburne M., Selitrenkoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Koth G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamvysseis M.,
RA Mauceli E., Bielke C., Rudd S., Frishman D., Krystofova S.,
RA Rasmussen C., Metznerberg R.J., Perkins D.D., Kroken S., Cogoni C.,
RA Macino G., Catchside D., Li W., Pratt R.J., Omani S.A.,
RA Desouza C.P., Glass L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nuebaum C., Birren B.;
RT "The genome sequence of the filamentous fungus Neurospora crassa."
RL Nature 422:859-868(2003).
CC -!- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
CC diphosphate + L-leucyl-tRNA(Leu).
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M30472; AAA33599.1; --
CC EMBL; AA451018; CAC18253.1; --
CC EMBL; AABX01000273; EAA31967.1; --
CC PIR; A33474; SYNCLM.
CC
CC InterPro; IPR002302; Leu-tRNA synthetase.
CC InterPro; IPR002300; tRNA-synt 1a.
CC InterPro; IPR001412; tRNA-synt 1.
CC InterPro; IPR009080; tRNA synthetase.
CC InterPro; IPR009008; tRNA synthetase.
CC Pfam; PF00133; tRNA-synt 1; 1.
CC PRINTS; PR00985; TRNASYNTHLEU.
CC TIGRFam; TIGR00396; leuS bact. 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE I; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Ligase; Mitochondrion;
KW Protein biosynthesis; Translational peptide.
FT TRANSIT 1 52 Mitochondrion.
FT CHAIN 53 994 Leucyl-tRNA synthetase.
FT SITE 93 103 "HIGH" region.
FT SITE 697 701 "KMSKS" region.
FT BINDING 700 700 ATP (by similarity).
FT SEQUENCE 994 AA; 109853 MW; 2ADB26302A8465F9 CRC64;
SQ
Query Match 48.3%; Score 42; DB 1; Length 994;
Best Local Similarity 60.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 5 HDLRAYGFWR 14
Db 403 HDLRDHAFWK 412

```

```

RESULT 28
Q9NEG4 PRELIMINARY; PRT; 306 AA.
ID Q9NEG4
AC Q9NEG4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein L2954.12.
GN Name=L2954.12;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Borzym K., Klages S., Reinhardt R., Beck A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138618; CAB71284.2; -.
KW Hypothetical protein.
SQ SEQUENCE 306 AA; 34129 MW; D03DA12477200AE7 CRC64;
Query Match 47.7%; Score 41.5; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 0; Indels 3; Gaps 1;
Qy 4 QHDLRAYGFWR 15
Db 213 QHN---YGFWR 221
RESULT 29
Q7Z1X3 PRELIMINARY; PRT; 60 AA.
ID Q7Z1X3
AC Q7Z1X3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2-macroglobulin family (Fragment).
OS Botryllus schlosseri (Star ascidian).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Botryllidae; Botryllus.
OX NCBI_TaxID=30301;
RN [1]
RP SEQUENCE FROM N.A.
RA Bartl S., Balsh M., Weissman I.L., Diaz M.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY102167; AAM52096.1; -.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR InterPro; IPR008930; terp_cyc_toroid.
FT NON TER 1 60
FT NON TER 60 60
SQ SEQUENCE 60 AA; 6911 MW; 9637A853A698D685 CRC64;
Query Match 47.1%; Score 41; DB 2; Length 60;
Best Local Similarity 46.2%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy 1 RSYQHLRAYGFWR 13
Db 39 QNYRHDNGAYSIM 51
RESULT 30
Q9H356

```

```
ID Q9H356 PRELIMINARY; PRT; 61 AA.
AC Q9H356;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PRO1804.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver.
RA Zhang C., Yu Y., Zhang S., Zhou G., Wei H., Bi J., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF132201; AAC35547.1; -.
SQ SEQUENCE 61 AA; 6678 MW; F238B05082E080C9 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 61;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDLRAYGFW 13
Db 6 SYQNSLLGLGFW 17

RESULT 31
Q8D683 PRELIMINARY; PRT; 61 AA.
AC Q8D683;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transposase.
GN OrderedLocusNames=VV20654;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016810; AAC07596.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00665; rve; 1.
KW Complete proteome.
SQ SEQUENCE 61 AA; 7115 MW; A424640F25849FA7 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 61;
Best Local Similarity 46.2%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFW 13
Db 13 RQYRKQLQGYGMW 25

RESULT 32
Q9KT60 PRELIMINARY; PRT; 188 AA.
AC Q9KT60;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RNA polymerase sigma-70 factor, ECF subfamily.

ID Q9H356 OrderedLocusNames=VC1045;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004186; AAP94204.1; -.
DR FIR; D82248; D82248.
DR TIGR; VC1045; -.
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR000838; Sigma70_ECF.
DR InterPro; IPR007627; Sigma70_r2.
DR Pfam; PF04542; Sigma70_r2; 1.
DR PROSITE; PS01063; SIGMA70_ECF; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 188 AA; 22547 MW; 3313FA82010855A2 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 188;
Best Local Similarity 46.2%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFW 13
Db 32 RAYHRDLRYAYW 44

RESULT 33
Q7MIS3 PRELIMINARY; PRT; 191 AA.
AC Q7MIS3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RNA polymerase sigma-70 factor, ECF subfamily.
GN OrderedLocusNames=VV2442;
OS Vibrio vulnificus (strain VU016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AF005339; BAC95206.1; -.
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR000838; Sigma70_ECF.
DR InterPro; IPR007627; Sigma70_r2.
DR Pfam; PF04542; Sigma70_r2; 1.
DR PROSITE; PS01063; SIGMA70_ECF; UNKNOWN_1.
```

KW Complete proteome.
SQ SEQUENCE 191 AA; 22641 MW; 194A1059BF0B0C89 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 191;
Best Local Similarity 46.2%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 35 RAYHRDLRYAYW 47

RESULT 34
Q8DB49 PRELIMINARY; PRT; 191 AA.
AC Q8DB49;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RNA polymerase sigma-70 factor.
GN OrderedLocusNames=VV11974;
OS *Vibrio vulnificus*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of *Vibrio vulnificus* CMCP6";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016803; AA010374.1; -
DR GO; GO:0016987; F:sigma factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006352; P:transcription initiation; IEA.
DR InterPro; IPR000838; Sigma70_BCF.
DR Pfam; PF04542; Sigma70_x2; 1.
DR PROSITE; PS01063; SIGMA70_BCF; UNKNOWN1.
KW Complete proteome.
SQ SEQUENCE 191 AA; 22641 MW; 194A1059BF0B0C89 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 191;
Best Local Similarity 46.2%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 35 RAYHRDLRYAYW 47

RESULT 35
Q8PW04 PRELIMINARY; PRT; 229 AA.
AC Q8PW04;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transporter.
GN OrderedLocusNames=MM1800;
OS *Methanosarcina mazei* (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goe1 / Goe1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierer A., Baumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,

RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of *Methanosarcina mazei*: evidence for lateral gene
transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013416; AAM31496.1; -
DR InterPro; IPR003744; DUF165.
DR Pfam; PF02592; DUF165; 1.
DR TIGRFAMs; TIGR00697; DUF165; 1.
KW Complete proteome.
SQ SEQUENCE 229 AA; 25476 MW; D5495B8192985291 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 229;
Best Local Similarity 46.2%; Pred. No. 89;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGFW 14
Db 110 SYQHDRLRAYGFW 122

RESULT 36
Q8G977 PRELIMINARY; PRT; 260 AA.
AC Q8G977;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase (fragment).
GN Name=tnpA;
OS *Pseudomonas putida*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PP3;
RX MEDLINE=22313464; PubMed=12426347;
RX DOI=10.1128/JB.184.23.6581-6591.2002;
RA Weightman A.J., Topping A.W., Hill K.E., Lee L.;
RT "Investigation of two evolutionarily unrelated halocarboxylic acid
dehalogenase gene families";
RL J. Bacteriol. 184:6581-6591(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PP3;
RA Hill H.K., Weightman A.J.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ534881; CAD59132.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR02560; Transposase_12.
DR Pfam; PF01610; Transposase_12; 1.
DR NON_TER 1
FT NON_TER 260
SQ SEQUENCE 260 AA; 30910 MW; 32BC1FA00D6ABF3B CRC64;

Query Match 47.1%; Score 41; DB 2; Length 260;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGFW 13
Db 228 AYEHKRFYGIW 239

RESULT 37
Q8GBV1 PRELIMINARY; PRT; 260 AA.
AC Q8GBV1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase (fragment).
GN Name=tnpA;
OS Pseudomonas sp. F12.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=217730;
[1]
RN PSEUDOMONAS SP. F12
RP SEQUENCE FROM N.A.
RC STRAIN=F12;
RA Hill K.E., Weightman A.J.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ534880; CAD59131.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR002560; Transposase 12.
DR Pfam; PF01610; Transposase_12; 1.
FT NON_TER 1
FT NON_TER 260
SQ SEQUENCE 260 AA; 30910 MW; 32BC1FA00D6ABF3B CRC64;

Query Match 47.1%; Score 41; DB 2; Length 260;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGF 13
Db 228 AYEHEKRFYGIW 239
:|:|:|:|

RESULT 38
RLS_NEUCR
ID - RL5_NEUCR STANDARD; PRT; 301 AA.
AC OS9553; Q7RVB8;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 23-OCT-2004 (Rel. 45, Last annotation update)
DE 60S ribosomal protein L5 (CP4).
GN Name=rpl-5; Synonyms=crp-4; ORFNames=B7H23.350, NCU04331.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
[1]
RN - SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RA de la Serna I.L., Cujec T.P., Shi Y., Tyler B.M.;
RT "Non-coordinate regulation of 5S rRNA genes and the gene encoding the
RT 5S rRNA-binding ribosomal protein homolog in Neurospora crassa.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RA MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gkg293;
RA Mannhaupf G., Montrone C., Haase D., Mewes H.-W., Algn V.,
RA Hohelsel J.D., Partmann B., Nyakatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence.";
RL Nucleic Acids Res. 31:1944-1954 (2003).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianklief P., Bell-Pedersen D., Nelson M.A.,
RA Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamvysselis M.,
RA Mauceli E., Bielke C., Rudd S., Frishman D., Krystofova S.,
RA Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S., Cogoni C.,
RA Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.P., Glass L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 422:859-868 (2003).
CC -!- FUNCTION: This protein binds 5S rRNA (by similarity).
CC -!- SIMILARITY: Belongs to the ribosomal protein L18p family.
CC -!- SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF054907; AAC09000.1; -.
DR EMBL; BX294026; CAD71058.1; -.
DR EMBL; AABX01000314; EAA31342.1; -.
DR InterPro; IPR005484; Ribosomal_L18p.
DR InterPro; IPR005485; Ribosomal_L5euk.
DR Pfam; PF00861; Ribosomal_L18p; 1.
DR PRINTS; PR00058; RIBOSOMALL5.
DR ProDom; PD001394; Ribosomal_L18p; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 301 AA; 34411 MW; 289B0F0D1672A5F9 CRC64;

Query Match 47.1%; Score 41; DB 1; Length 301;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYG 11
Db 78 AYSHELKAYG 87
:|:|:|:|

RESULT 39
Q6MAL4
ID Q6MAL4 PRELIMINARY; PRT; 323 AA.
AC Q6MAL4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=pc1661;
OS Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.). Parachlamydia.
OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiales; Parachlamydiaceae; Parachlamydia.
OX NCBI_TaxID=264201;
[1]
RN SEQUENCE FROM N.A.
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA Partmann B., Brandt P., Nyakatura G.J., Droese M., Frishman D.,
RA Rattei T., Mewes H.-W., Wagner M.;
RT "Genome sequence of an amoeba symbiont and its use for reconstructing
RT the evolutionary history of chlamydiae.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX908798; CAF24385.1; -.
DR InterPro; IPR006597; Sel_like.
DR SMART; SM00671; SEL1; 5.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 323 AA; 36644 MW; 0519D5135B9F8B08 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 323;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGF 15
Db 308 SYENRLRIYHWN 321
:|:|:|:|:
```

RESULT 40
Q9Z5V9
ID Q9Z5V9 PRELIMINARY; PRT; 338 AA.
AC Q9Z5V9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase A.
GN Name=tnpA;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=142;
RX MEDLINE=92240444; PubMed=10224014;
RA Tsoi T.V., Plotnikova E.G., Cole J.R., Guerin W.F., Bagdasarian M.,
RA Tiedje J.M.;
RT "Cloning, expression, and nucleotide sequence of the Pseudomonas
RT aeruginosa 142 ohb genes coding for oxygenolytic ortho dehalogenation
RT of halobenzoates";
RL Appl. Environ. Microbiol. 65:2151-2162(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=142;
RA Hrywna Y., Tsoi T.V., Maltseva O.V., Quensen III J.F., Tiedje J.M.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121970; AAD20008.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR002560; Transposase_12.
DR Pfam; PF01610; Transposase_12; 1_
SQ SEQUENCE 338 AA; 39711 MW; E0C3DF2EABE3B329 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 338;
Best Local Similarity 50.0%; Pred. NO. 1.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 13
Db :|||
202 AYEHKERYGIW 213

Search completed: September 16, 2005, 12:43:53
Job time : 173 secs

THIS PAGE BLANK (user)